

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:02:42 ; Search time 2891.94 Seconds
(without alignments)
10100.295 Million cell updates/sec

Title: US-09-835-992A-19

Perfect score: 714

Sequence: 1 cgccagaaaagtcatttta.....caccgtgttgcacattca 714

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sfs:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sfs:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	710	99.4	714	6	ARI146579	ARI146579 Sequence
2	710	99.4	714	6	BD079828	BD079828 Cancer-as
3	672.8	94.2	1052	9	HSSCPXT16	U11313 Human stero
4	672.8	94.2	2572	9	HUMSCP2A	M75883 Human stero
5	672.8	94.2	122176	2	AL358233	AL358233 Homo sapi
6	672.8	94.2	175046	2	AC022728	AC022728 Homo sapi
7	672.8	94.2	193774	9	AL445183	AL445183 Human DNA
8	667.6	93.5	1439	9	BC005911	BC005911 Homo sapi
9	611.8	85.7	687	6	ARI146580	ARI146580 Sequence
10	611.8	85.7	687	6	BD079829	BD079829 Cancer-as
11	474.6	66.5	1229	9	S52450	S52450 sterol carr
12	459.8	64.4	1219	9	HUMSTEA	M55421 Human stero
13	436.2	61.1	843	6	ARI146582	ARI146582 Sequence
14	436.2	61.1	843	6	BD079831	BD079831 Cancer-as
15	318.4	44.6	1500	9	HUMSCP2B	M75884 Human stero
16	250.6	35.1	2661	4	AF051897	AF051897 Oryctolag
17	202.8	28.4	2599	10	RAT60KDA	M62763 Rat 60 kDa
18	194.6	27.3	145564	2	AC134794	AC134794 Mus muscu
19	194.4	27.2	2571	10	RATSCPXA	M57453 Rat sterol
20	193.2	27.1	2626	10	BC018384	BC018384 Mus muscu
21	193.2	27.1	248677	2	AL844206	AL844206 Mus muscu
22	169.8	23.8	904	10	S80339	S80339 SCP-2=stero
23	165.6	23.2	173	6	AR246194	AR246194 Sequence
24	107.4	15.0	263	6	A74403	A74403 Sequence 89
25	107.4	15.0	263	6	A77382	A77382 Sequence 89
26	89.2	12.5	151259	2	AC138822	AC138822 Homo sapi
27	89.2	12.5	160701	9	AC108106	AC108106 Homo sapi
28	89.2	12.5	172404	9	AC138948	AC138948 Homo sapi
29	89.2	12.5	181154	2	AC138842	AC138842 Homo sapi
30	89.2	12.5	185785	9	AC138817	AC138817 Homo sapi
31	89.2	12.5	188885	9	AC138850	AC138850 Homo sapi
32	89.2	12.5	192925	2	AC138953	AC138953 Homo sapi
33	89.2	12.5	201419	9	AC138908	AC138908 Homo sapi
34	89.2	12.5	211297	2	AC138860	AC138860 Homo sapi
35	89.2	12.5	216200	2	AC138862	AC138862 Homo sapi
36	89.2	12.5	225384	9	AC138832	AC138832 Homo sapi
37	89.2	12.5	238249	2	AC138960	AC138960 Homo sapi
38	87.6	12.3	130416	9	AC008851	AC008851 Homo sapi
39	85.2	11.9	113202	9	HSJ397H23	AL121972 Human DNA
40	79.8	11.2	149597	2	AC034271	AC034271 Homo sapi
41	79.8	11.2	157749	2	AC025535	AC025535 Homo sapi
42	78.8	11.0	273275	3	AE014828	AE014828 Plasmodiu
43	77.8	10.9	8056	6	AX599046	AX599046 Sequence
44	77.4	10.8	125623	3	AC115599	AC115599 Dictyoste
45	77.2	10.8	37957	8	U17009	U17009 Phytosphor

ALIGNMENTS

RESULT 1
ARI146579
LOCUS ARI146579 714 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 19 from patent US 6218521.
ACCESSION ARI146579
VERSION ARI146579.1 GI:15109768
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 714)
AUTHORS Obata,Y.
TITLE Isolated nucleic acid molecules associated with gastric cancer and
JOURNAL Patent: US 6218521-A 19 17-APR-2001;

FEATURES		Location/Qualifiers	
source		1..714	
		/Organism="unknown"	
BASE COUNT	258 a	100 c	92 g 260 t 4 others
ORIGIN			
Query Match 99.4%; Score 710; DB 6; Length 714;			
Best Local Similarity 100.0%; Pred. No. 1.5e-93;			
Matches 714; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGCCAGAAAAAGTATTTTAATTTCTATTAACATCTCTCAAGCATTATTATCC	60
Db	1	CGCCAGAAAAAGTATTTTAATTTCTATTAACATCTCTCAAGCATTATTATCC	60
QY	61	TATATCTCACTGAATTTTAAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATAA	120
Db	61	TATATCTCACTGAATTTTAAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATAA	120
QY	121	ATGCAGATAATTAACCTTACATGAAAAAGGAAAAATTATACAAGAGCTGAGAACGTTAT	180
Db	121	ATGCAGATAATTAACCTTACATGAAAAAGGAAAAATTATACAAGAGCTGAGAACGTTAT	180
QY	181	AAATGGAATGAGATTATAAATTTGAAAACTGCATCTGAAGCAAACTTATTGTTCAATT	240
Db	181	AAATGGAATGAGATTATAAATTTGAAAACTGCATCTGAAGCAAACTTATTGTTCAATT	240
QY	241	ATNCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT	300
Db	241	ATNCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT	300
QY	301	AAAAATATTTTATTTTAAATAAAGCCCTGTGTCAAGCTCTGATCATATTCTTTATT	360
Db	301	AAAAATATTTTATTTTAAATAAAGCCCTGTGTCAAGCTCTGATCATATTCTTTATT	360
QY	361	TTGATTGGGAANAAAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTAGATTTTTA	420
Db	361	TTGATTGGGAANAAAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTAGATTTTTA	420
QY	421	ATCTCACTAATTTTAAANAACATAATGAGAAATTGATTAATGACATGAAGTGACAACTA	480
Db	421	ATCTCACTAATTTTAAANAACATAATGAGAAATTGATTAATGACATGAAGTGACAACTA	480
QY	481	ATTACTGGCCAGCTGTGGCACTGTGTTCTTACTTAGTCTCCCAAGGAAACTCTTA	540
Db	481	ATTACTGGCCAGCTGTGGCACTGTGTTCTTACTTAGTCTCCCAAGGAAACTCTTA	540
QY	541	AATTGAATCTTCAGCAGATAATCCTTAATATACCTTTGTAAGCAAAACAAAGCTTTT	600
Db	541	AATTGAATCTTCAGCAGATAATCCTTAATATACCTTTGTAAGCAAAACAAAGCTTTT	600
QY	601	TGTTTACATAGTCTCTTGGGATTTTACTGTCTTAATTTTATCTGAAACTCAATTTTAC	660
Db	601	TGTTTACATAGTCTCTTGGGATTTTACTGTCTTAATTTTATCTGAAACTCAATTTTAC	660
QY	661	CCCAGACCATAATTAACCATATTAATCTGTGTTNGACAGAGTTGTTGCCAATTCA	714
Db	661	CCCAGACCATAATTAACCATATTAATCTGTGTTNGACAGAGTTGTTGCCAATTCA	714
RESULT 2			
LOCUS	BD079828	714 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Cancer-associated nucleic acids and polypeptides.		
ACCESSION	BD079828		
VERSION	BD079828.1	GI:22625431	
KEYWORDS	JP 2001516009-A/494.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 714)		
AUTHORS	Oghare, M., Obata, Y., Pfreundschuh, M., Tureci, O. and Sahin, U.		

TITLE	Cancer-associated nucleic acids and polypeptides									
JOURNAL	Patent: JP 2001516009-A 494 25-SEP-2001;									
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH									
	OS	Homo sapiens (human)								
	PN	JP 2001516009-A/494								
	PD	25-SEP-2001								
	PF	15-JUL-1998 JP 2000503425								
	PR	17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR								
	10-OCT-1997 US	60/061765,10-OCT-1997 US 08/948705 PR								
	11-OCT-1997 GB	9721697.2,22-JUN-1998 US 09/102322 PI LLOYD								
	CHEN,	J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG								
	PI	IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI								
	PI	OZLEM TURECI,								
	PI	UGUR SAHIN								
	PC	G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC								
		A61P35/00,								
	PC	C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,								
	PC	C12N15/00								
	CC	Cancer-associated nucleic acids and polypeptides. FH Key								
		Location/Qualifiers								
	FT	source 1..714								
	FT	Location/Qualifiers /organism='Homo sapiens (human)'. 1..714								
FEATURES	source	1..714								
		/organism="Homo sapiens"								
		/mol_type="genomic DNA"								
		/db_xref="taxon:9606"								
BASE COUNT	258 a	100 c	92 g	260 t	4 others					
ORIGIN										
Query Match	99.4%;	Score 710;	DB 6;	Length 714;						
Best Local Similarity	100.0%;	Pred. No. 1.5e-93;								
Matches 714;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;						
QY	1	CGCCAGAAAAAGTATTTTAATTTTCTATTAACATCTCTCAAGCATTATTTATCC	60							
Db	1	CGCCAGAAAAAGTATTTTAATTTTCTATTAACATCTCTCAAGCATTATTTATCC	60							
QY	61	TATATCTCACTGAATTTTAAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATTA	120							
Db	61	TATATCTCACTGAATTTTAAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATTA	120							
QY	121	ATGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAGAGACTGAGAACGTTAT	180							
Db	121	ATGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAGAGACTGAGAACGTTAT	180							
QY	181	AAATGGAATGAGATTATAATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATT	240							
Db	181	AAATGGAATGAGATTATAATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATT	240							
QY	241	ATNCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT	300							
Db	241	ATNCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT	300							
QY	301	AAAAATATTTTATTTTAAANAATAGCCCTGTGTTCAAGCTCTGATCATATTCTTTATT	360							
Db	301	AAAAATATTTTATTTTAAANAATAGCCCTGTGTTCAAGCTCTGATCATATTCTTTATT	360							
QY	361	TTGATTGGGAANAAAAATACTGTTCTGATAGCATGAAATGCAAAATTTTAGATTTTTA	420							
Db	361	TTGATTGGGAANAAAAATACTGTTCTGATAGCATGAAATGCAAAATTTTAGATTTTTA	420							
QY	421	ATCTCACTAATTTTAAANAACATAATGAGAAATTGATTAATGACATGAAGTGACAACTA	480							
Db	421	ATCTCACTAATTTTAAANAACATAATGAGAAATTGATTAATGACATGAAGTGACAACTA	480							
QY	481	ATTACTGGCCAGCTGTGGCATGTGTTTCTTACTTAGTCTCCCAAGGAAACTCTTA	540							
Db	481	ATTACTGGCCAGCTGTGGCATGTGTTTCTTACTTAGTCTCCCAAGGAAACTCTTA	540							
QY	541	AATTGAATCTTCAGCAGATAATCCTTAATATATACTTTGTAAGCAAAACAAAGCTTTT	600							

|||||
Db 541 AATTGAATCTTCAGACAGATATATCTTAATATACCTTGTAAACCAAAAGCTTTT 600
QY 601 TGTTTACATAGTCTTGGGATTTTACTGTCTCTAATTTTATCTGAAACTCAATTTTAC 660
Db 601 TGTTTACATAGTCTTGGGATTTTACTGTCTCTAATTTTATCTGAAACTCAATTTTAC 660
QY 661 CCCAGACCATAATTACCATATTAACCTTGTNTGCACAGTGTGGCCAATTCA 714
Db 661 CCCAGACCATAATTACCATATTAACCTTGTNTGCACAGTGTGGCCAATTCA 714
RESULT 3
HSSCPXT16/c 1052 bp DNA linear PRI 25-JAN-1997
LOCUS HSSCPXT16 1052 bp DNA linear PRI 25-JAN-1997
DEFINITION Human sterol carrier protein-X/sterol carrier protein-2
(SCP-X/SCP-2) gene, exon 16, and complete cds.
ACCESSION U11313
VERSION U11313.1 GI:532077
KEYWORDS 16 of 16
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1052)
AUTHORS Ohbe,T., Rennert,H., Pfeiffer,S.M., He,Z., Yamamoto,R., Holt,J.A.,
Billheimer,J.T. and Straus,J.F. III.
TITLE The structure of the human sterol carrier protein X/sterol carrier
protein 2 gene (SCP2)
JOURNAL Genomics 24 (2), 370-374 (1994)
MEDLINE 95213031
PUBMED 7698762
REFERENCE 2 (bases 1 to 1052)
AUTHORS Straus,J.F. III.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1994) Jerome F. Straus III, Department of
Obstetrics and Gynecology, Division of Reproductive Biology,
University of Pennsylvania Medical Center, 778 Clinical Research
Building, 422 Curie Boulevard, Philadelphia, PA 19104-6142, USA
FEATURES
Source Location/Qualifiers
1..1052
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p32"
/cfeature_type="11ver"
join(U11297..1..3666,U11299..1..108,U11300..1..122,
U11301..1..182,U11302..1..115,U11303..1..177,
U11304..1..114,U11305..1..137,U11306..1..201,
U11307..1..198,U11308..1..158,U11309..1..1725,
U11310..1..153,U11311..1..180,U11312..1..130,1..1034)
/partial
/gene="SCP-X/SCP-2"
join(U11297..1:3452..3641,U11299..1:26..83,U11300..1:26..97,
U11301..1:26..157,U11302..1:26..90,U11303..1:26..152,
U11304..1:26..89,U11305..1:26..112,U11306..1:26..176,
U11307..1:26..173,U11308..1:26..133,U11309..1:1547..1700,
U11310..1:26..128,U11311..1:26..155,U11312..1:26..105,
26..1034)
/gene="SCP-X/SCP-2"
/product="sterol carrier protein-X/sterol carrier
protein-2"
join(U11297..1:3573..3641,U11299..1:26..83,U11300..1:26..97,
U11301..1:26..157,U11302..1:26..90,U11303..1:26..152,
U11304..1:26..89,U11305..1:26..112,U11306..1:26..176,
U11307..1:26..173,U11308..1:26..133,U11309..1:1547..1700,
U11310..1:26..128,U11311..1:26..155,U11312..1:26..105,
26..121)
/gene="SCP-X/SCP-2"
/codon_start=1
/product="sterol carrier protein-X/sterol carrier

protein-2"
/protein_id="AAB41286.1"
/db_xref="GI:1773239"
/translation="MSSSPWEPAALRVFVVGVMKTFVKPGAEASRDYFDLAEAGK
KALADQIPYSAVDQACVGVFSDTCGQRAIYHSLGMTGPIINVNNGATGSTALL
MARQLDGVAAECVLALGFEMSKSLGIFSDRTIPYDKHVDLIRKYLGAHPVAP
QMFYAGKEHMEKYGKTIHFPAIKGNHKSVMNPSQFQDEYSLDEVMASKEVDF
LTLLCCPTSDGAAAILASEAFQKYLQSKAVEILAQEMMTDLPSSFEKSIIMV
GFDMSKEARKCYEKSGLTPNDIDVIELHDCFSTNELTYEALGLCPGQATLVDRG
DNTYGGKVINPSGGLISKGPLGATGLAQCAELCMQLRGEGAKOVGAKVALQHLN
DIGAVVTVLLYKMGFPPEAASSFRTHQIEAVPTSSASDGFKANLYEKEIEKLEEGEO
FVKKIGGIFAFKVKDGPGEKEATVVDVKNKGSVLPSNDDKADCTITMADSDFLALM
TGKMPQSAFFQGLKITGNMGLMKLQNLQLPGNAKL"
26..1034
/gene="SCP-X/SCP-2"
/note="nucleotides 122-1027 are not translated"
/number=16
BASE COUNT 352 a 157 c 164 g 379 t
ORIGIN
Query Match 94.2%; Score 672.8; DB 9; Length 1052;
Best Local Similarity 98.2%; Pred. No. 3e-88;
Matches 699; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 2 GCCAGAAAAAGTTATTTTAAATTTCTATTAACAATCTTCTCAAGCATTTATTCCT 61
Db 1027 GCCAGAAAAAGTTATTTTAAATTTCTATTAACAATCTTCTCAAGCATTTATTCCT 968
QY 62 ATATCTCACTGAATTTTAAAGAAATTAACATTAAGTATTAAGAAAACTAGAAAAAGATAAA 121
Db 967 ATATCTCACTGAATTTTAAAGAAATTAACATTAAGTATTAAGAAAACTAGAAAAAGATAAA 908
QY 122 TGCAGATAATTAACCTTACATGAAGAAAGAAATTAACAAGAAAGACTGAGACGTATA 181
Db 907 TGCAGATAATTAACCTTACATGAAGAAAGAAATTAACAAGAAAGACTGAGACGTATA 848
QY 182 AATTGAATGAGATTATTAATTTGAAGAACTGCATCTGAAGCAAACTTATTTCAATTA 241
Db 847 AATTGAATGAGATTATTAATTTGAAGAACTGCATCTGAAGCAAACTTATTTCAATTA 788
QY 242 TNCCTAATGATGGTGTATTGACTAATACACTGATTTTCAAGAGAAAGCAATGTTA 301
Db 787 TTCTTAATGATGGTGTATTGACTAATACACTGATTTTCAATTAAGAAAGCAATGTTA 728
QY 302 AAAATATTTTATTTTAAATAATTAAGCCGTGTGTCAGAGCTCTGATCATATTTCTTTATTT 361
Db 727 AAAATATTTTATTTTAAATAATTAAGCCGTGTGTCAGAGCTCTGATCATATTTCTTTATTT 668
QY 362 TGATTTGGAAANAATACTGTTCTGATAGCATGAAGAAATGCAAAATTTTGAATTTTAA 421
Db 667 TGATTTGGAAANAATACTGTTCTGATAGCATGAAGAAATGCAAAATTTTGAATTTTAA 608
QY 422 TCTCACTAATTTTAAANAATACTATTGAGAAATGATTAATGACATGAGTGCACACACTAA 481
Db 607 TCTCACTAATTTTAAAGAACTATTGAGAAATGATTAATGACATGAGTGCACACACTAA 548
QY 482 TTACTGGCCAGCTGTGGCATTTGTGTTCTTAATTAAGTTCTCCCAAGGGAACCTCTTAA 541
Db 547 TTACTGGCCAGCTGTGGCATTTGTGTTCTTAATTAAGTTCTCCCAAGGGAACCTCTTAA 489
QY 542 ATGAATCTTCAGCAGATAATCTCTTAATAATTAATCTTGTGAAGCAAAACAAAGCTTTT 601
Db 488 ACTGAATCTTCAGCAGATAATCTCTTAATAATTAATCTTGTGAAGCAAAACAAAGCTTTT 429
QY 602 GTTTACATAGTCTTGGGATTTTACTGTCTAATTTTATTTGAAACTCAATTTTACC 661
Db 428 GTTTACATAGTCTTGGGATTTTACTGTCTAATTTTATTTGAAACTCAATTTTACC 370
QY 662 CCAGACCATAATTAACATATTAATCTTGTNTGCACAGTGTGGCAATTC 713
Db 369 CCAGACCATAATTAACATATTAATCTTGTNTGCACAGTGTGGCAATTC 318

LOCUS DEFINITION	2572 bp	mRNA	linear	PRI 06-DEC-1993
LOCUS HUMSCP2A/c				
DEFINITION	Human sterol carrier protein X/sterol carrier protein 2 mRNA, complete cds.			
ACCESSION	M75883			
VERSION	M75883.1	GI:432974		
KEYWORDS	sterol carrier protein-2, sterol carrier protein X.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 2572)			
TITLE	He, Z., Yamamoto, R., Furth, E.E., Schantz, L.J., Naylor, S.L., George, H., Billheimer, J.T. and Strauss, J.F. III. CDNAs encoding members of a family of proteins related to human sterol carrier protein 2 and assignment of the gene to human chromosome 1 p21---pter			
JOURNAL	DNA Cell Biol. 10 (8), 559-569 (1991)			
MEDLINE	92029618			
PUBMED	1718316			
REFERENCE	2 (bases 1 to 2572)			
AUTHORS	Vesa, J., Hellsten, E., Branoski, B.L., Emanuel, B.S., Billheimer, J.T., Mead, S., Cowell, J.K., Strauss, J.F. III, and Peltonen, L. Assignment of sterol carrier protein X/sterol carrier protein 2 to 1p32 and exclusion as the causative gene for infantile neuronal ceroid lipofusiosis			
TITLE	Unpublished			
JOURNAL	On Dec 6, 1993 this sequence version replaced gi:410029.			
COMMENT	Original source text: Human liver cDNA to mRNA.			
FEATURES	location/Qualifiers			
source	1. .2572			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/map="1p32"			
	/tissue_type="liver"			
	/clone_1ib="Clontech HL-1001b,HL-1115b"			
gene	1. .2572			
	/gene="SCP-X/SCP-2"			
CDS	22. .1665			
	/gene="SCP-X/SCP-2"			
	/codon_start=1			
	/product="sterol carrier protein-X/sterol carrier protein-2"			
	/protein_id="AAA03557.1"			
	/db_xref="GI:432975"			
	/translation="MSSSPWEPATLRVPVGVGMTKFKVPGAEENSRDYPDLAEAGK KALADAOIPYSAVDQACVGVFGDSTCGRAIYHSLGMTGPIINNNNCATGSTALF MARQLIOGGVAECVLAIGFEKMSKSLGIKFSDRITPTDKHVDLLINKYGLSAHPVAP QMFYAGKEHMEKYGTKIEHFAKIGWKNHGSVNNPYSQFDEYSIDETMASKEVEFDE LTILQCCPTSDGAAAILASAEFAVQKYGLOSKAVEILAQEMMTDLPSEFEKSIKIMV GFDMSCKEARKCYEKSGLTPNDIDVIELHDCFSITNELLTYEALGLCPREGOGATLVDRG DNTYGGKWINPSGGLISKGHPGLGATGLAOCALCWOLRGEAGKROYGAKVALQHNL GIGGAVVVTLYKMGFPEAASSFRTHQIEAVPTSSASDGFKANLVFKELEKLEEGEQ FVKKIGGIFAFKYKDGPGGKEATWVVDVKNKGSVLPSNDSKADCTITMADSDFLAM TGRKNPQSAFFQGLKITGMGLAMKLQNLQDPGNAKL"			
	1932. .1937			
polyA_signal	/gene="SCP-X/SCP-2"			
	/evidence=not_experimental			
polyA_signal	2213. .2218			
	/gene="SCP-X/SCP-2"			
	/evidence=not_experimental			
polyA_signal	2258. .2263			
	/gene="SCP-X/SCP-2"			
	/evidence=not_experimental			
polyA_signal	2341. .2346			
	/gene="SCP-X/SCP-2"			
	/evidence=experimental			
polyA_signal	2550. .2555			
	/gene="SCP-X/SCP-2"			
	/evidence=experimental			
polyA site	2572			

BASE COUNT	794 a	439 c	570 g	769 t
ORIGIN	/gene="SCP-X/SCP-2" /evidence=experimental			
Query Match	94.2%; Score 672.8; DB 9; Length 2572;			
Best Local Similarity	98.2%; Pred. No. 2:2e-88;			
Matches 699; Conservative	0; Mismatches 11; Indels 2; Gaps 2;			
QY	2	GCCAGAAAAAGTTATTTTAAATTTTCTATTAACAATCTCTCTCAAGCATTTATTTATCCT	61	
DB	2571	GCCAGAAAAAGTTATTTTAAATTTTCTATTAACAATCTCTCTCAAGCATTTATTTATCCT	2512	
QY	62	ATATCTCAGCTGAATTTTAAAGAAATACATTAGTATTAGAAAAAATAAGAAAAAGATATA	121	
DB	2511	ATATCTCAGCTGAATTTTAAAGAAATACATTAGTATTAGAAAAAATAAGAAAAAGATATA	2452	
QY	122	TGCAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAAGGACTGAGAACGTTATA	181	
DB	2451	TGCAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAAGGACTGAGAACGTTATA	2392	
QY	182	AATTGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAAACTTTATTTGTTCAATTA	241	
DB	2391	AATTGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAAACTTTATTTGTTCAATTA	2332	
QY	242	TNCTTAATGATGGTGTCTTATGACTAATACATGATTTTTCAGAAGAAACCATGTTA	301	
DB	2331	TTCTTAATGATGGTGTCTTATGACTAATACATGATTTTTCAGAAGAAACCATGTTA	2272	
QY	302	AAAAATTTTATTTTAAATAAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAATT	361	
DB	2271	AAAAATTTTATTTTAAATAAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAATT	2212	
QY	362	TGATTTGGGAANAATACTCTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA	421	
DB	2211	TGATTTGGGAANAATACTCTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA	2152	
QY	422	TCTCACTAATTTTAAANAACATAATGAGAAATGATTAATGACATGAAGTGACACAATA	481	
DB	2151	TCTCACTAATTTTAAANAACATAATGAGAAATGATTAATGACATGAAGTGACACAATA	2092	
QY	482	TTACTGGCCAGCTGTGGCATTGTGTTCTTAACTTAGTCTCCCAAGGAAAACTCTTAA	541	
DB	2091	TTACTGGCCAGCTGTGGCATTGTGTTCTTAACTTAGTCTCCCAA-GGAAAACCTCTTAA	2033	
QY	542	ATTGAATCTTCAGACAGAATAATCCTTAATAATACTTTGTAAGCAAAACAAAGCTTTTT	601	
DB	2032	ATTGAATCTTCAGACAGAATAATCCTTAATAATACTTTGTAAGCAAAACAAAGCTTTTT	1973	
QY	602	GTTTACATAGTCTTTGGGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTTACC	661	
DB	1972	GTTTACATAGTCTTTT-GGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTTACC	1914	
QY	662	CCAGACCATAAATTACCATATTACTTTGTTNTGCACAGTTGTTGCCAATTTC	713	
DB	1913	CCAGACCATAAATTACCATATTACTTTGTTNTGCACAGTTGTTGCCAATTTC	1862	
RESULT 5				
AL358233	122176 bp DNA linear HTG 10-JUL-2001			
LOCUS	AL358233 Homo sapiens chromosome 1 clone RP5-835A17, *** SEQUENCING IN			
DEFINITION	PROGRESS ***, 8 unordered pieces.			
ACCESSION	AL358233			
VERSION	AL358233.3 GI:9797852			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_CANCELLED.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	1 McIay, K.			
	Direct Submission			

JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 12, 2000 this sequence version replaced gi:9214208.

COMMENT ----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: d0335A17

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator ET-amersham; 20% of reads Chemistry:

Dye-terminator Big Dye; 79% of reads

Consensus quality: 118880 bases at least Q40

Consensus quality: 120216 bases at least Q30

Consensus quality: 120979 bases at least Q20

Insert size: 121476; sum-of-contigs

Insert size: 136681; 7.7% error; agarose-fp

Quality coverage: 4.05x in Q20 bases; sum-of-contigs Quality

coverage: 3.75x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 13486: contig of 13486 bp in length
* 13487 13586: gap of 100 bp
* 13587 30893: contig of 17307 bp in length
* 30894 30993: gap of 100 bp
* 30994 37022: contig of 6029 bp in length
* 37023 37122: gap of 100 bp
* 37123 45381: contig of 8259 bp in length
* 45382 45481: gap of 100 bp
* 45482 56788: contig of 11307 bp in length
* 56789 56889: gap of 100 bp
* 56889 84391: contig of 27503 bp in length
* 84392 84491: gap of 100 bp
* 84492 117843: contig of 33352 bp in length
* 117844 117944: gap of 100 bp
* 117944 122176: contig of 4233 bp in length.

FEATURES Location/Qualifiers

source 1..122176

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP5-835A17"

/clone_lib="RPCT-5"

1..13486

/note="assembly fragment:01065

fragment_chain:1"

13587..30893

/note="assembly fragment:00181

fragment_chain:1"

30994..37022

/note="assembly fragment:00153

fragment_chain:1"

37123..45381

/note="assembly fragment:00580

fragment_chain:2"

45482..56788

/note="assembly fragment:00948

fragment_chain:2"

56889..84391

/note="assembly fragment:00192"

84492..117843

/note="assembly fragment:00299.0"
misc_feature 117944..122176
/note="assembly fragment:00926"
BASE COUNT 31142 a 30481 c 29651 g 30201 t 701 others
ORIGIN

Query Match 94.2%; Score 672.8; DB 2; Length 122176;
Best Local Similarity 98.2%; Pred. No. 5.6e-89;
Matches 699; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 2 GCCAGAAAAAGTATTATTTCTATTTAAACATTTCTCAAGCATTATTATTCCT 61

Db 47281 GCCAGAAAAAGTATTATTTCTATTTAAACATTTCTCAAGCATTATTATTCCT 47340

QY 62 ATATCTCACTGAATTTAAGAAATAACATTAGATTAGAAAACTAGAAAAAGATAAA 121

Db 47341 ATATCTCACTGAATTTAAGAAATAACATTAGATTAGAAAACTAGAAAAAGATAAA 47400

QY 122 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAAGACTGAGAACGTTATA 181

Db 47401 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAAGACTGAGAACGTTATA 47460

QY 182 AATTGAATGAGATTATAATTGAAAACCTGCATCTGAAGCAAACCTTATTGTTCAATTA 241

Db 47461 AATTGAATGAGATTATAATTGAAAACCTGCATCTGAAGCAAACCTTATTGTTCAATTA 47520

QY 242 TNCCTAATGATGGTGTGTTTATGACTAATACACTGATTTTTCAGAGAAACCACTGTTA 301

Db 47521 TNCCTAATGATGGTGTGTTTATGACTAATACACTGATTTTTCAGAGAAACCACTGTTA 47580

QY 302 AAAATATTTTATTTTAAATAAGCCGTGTGTTCAAGCTCTGATCATATTTCTTTATTT 361

Db 47581 AAAATATTTTATTTTAAATAAGCCGTGTGTTCAAGCTCTGATCATATTTCTTTATTT 47640

QY 362 TGATTTGGAAANAATACTGTTTCTGATAGCAATGCAAAATTTTATTTTAA 421

Db 47641 TGATTTGGAAANAATACTGTTTCTGATAGCAATGCAAAATTTTATTTTAA 47700

QY 422 TCTCACTAATTTAANAATACTATTGAGAAATGATTAATGACATGAGTGACACACTAA 481

Db 47701 TCTCACTAATTTAANAATACTATTGAGAAATGATTAATGACATGAGTGACACACTAA 47760

QY 482 TTACTGGCCAGCTGTGGCAATGTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAA 541

Db 47761 TTACTGGCCAGCTGTGGCAATGTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAA 47819

QY 542 ATGAATCTTCAGACAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTT 601

Db 47820 ATGAATCTTCAGACAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTT 47879

QY 602 GTTACATAGTCTTGGGATTTTACTGTTCCATAATTTATCTGAACCTCAATTTACC 661

Db 47880 GTTACATAGTCTTGGGATTTTACTGTTCCATAATTTATCTGAACCTCAATTTACC 47938

QY 662 CCAGACCATATTAACATTAATTAATTTGCAAGCTGTTGCCAATTC 713

Db 47939 CCAGACCATATTAACATTAATTAATTTGCAAGCTGTTGCCAATTC 47990

RESULT 6
AC022728/c AC022728 175046 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 1 clone RP11-310J14 map 1, WORKING DRAFT
DEFINITION
SEQUENCE, 34 unordered pieces.
ACCESSION AC022728 AC022728.4 GI:7249198
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175046)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 1, clone RP11-310J14
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175046)
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczký, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 175046)
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczký, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 16, 2000 this sequence version replaced gi:6980310. All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L5491

Center clone name: 310_J_14

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156135 bases at least Q40

Consensus quality: 165373 bases at least Q30

Consensus quality: 168980 bases at least Q20

Insert size: 185000; agarose-fp

Insert size: 171746; sum-of-contigs

Quality coverage: 3.4 in Q20 bases; agarose-fp

Quality coverage: 3.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 56: contig of 56 bp in length
57 156: gap of 100 bp
157 1191: contig of 1035 bp in length
1192 1291: gap of 100 bp
1292 2432: contig of 1141 bp in length
2433 2532: gap of 100 bp
2533 3742: contig of 1210 bp in length
3743 3842: gap of 100 bp
3843 4921: contig of 1079 bp in length
4922 5021: gap of 100 bp
5022 6073: contig of 1052 bp in length
6074 6173: gap of 100 bp
6174 7552: contig of 1379 bp in length
7553 7652: gap of 100 bp
7653 9001: contig of 1349 bp in length
9002 9101: gap of 100 bp
9102 10622: contig of 1521 bp in length
10623 10722: gap of 100 bp
10723 11836: contig of 1114 bp in length
11837 11936: gap of 100 bp
11937 13398: contig of 1462 bp in length
13399 13498: gap of 100 bp
13499 14949: contig of 1451 bp in length
14950 15050: gap of 100 bp
15050 16261: contig of 1211 bp in length
16261 16361: gap of 100 bp
16361 18265: contig of 1905 bp in length
18266 18365: gap of 100 bp
18366 19916: contig of 1551 bp in length
19917 20017: gap of 100 bp
20017 22939: contig of 2923 bp in length
22940 23039: gap of 100 bp
23040 26446: contig of 3406 bp in length
26446 26546: gap of 100 bp
26546 29327: contig of 2782 bp in length
29328 29427: gap of 100 bp
29428 32200: contig of 2773 bp in length
32201 32301: gap of 100 bp
32301 36738: contig of 4438 bp in length
36739 36838: gap of 100 bp
36839 43447: contig of 6609 bp in length
43448 43547: gap of 100 bp
43548 48525: contig of 4978 bp in length
48526 48625: gap of 100 bp
48626 52791: contig of 4166 bp in length
52792 52891: gap of 100 bp
52892 58391: contig of 5500 bp in length
58392 58491: gap of 100 bp
58492 64869: contig of 6378 bp in length
64870 64969: gap of 100 bp
64970 73369: contig of 8400 bp in length
73370 73469: gap of 100 bp
73470 81101: contig of 7632 bp in length
81102 81201: gap of 100 bp
81202 87754: contig of 6553 bp in length
87755 87854: gap of 100 bp
87855 95534: contig of 7680 bp in length
95535 95634: gap of 100 bp
95635 102797: contig of 7163 bp in length
102798 102897: gap of 100 bp
102898 111817: contig of 8920 bp in length
111818 111917: gap of 100 bp
111918 121160: contig of 9243 bp in length
121161 121260: gap of 100 bp
121261 140537: contig of 19277 bp in length
140538 140637: gap of 100 bp

```
FEATURES
Source
140638 175046: contig of 34409 bp in length.
Location/Qualifiers
1.175046
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-310J14"
/clone_1b="RPC1-11 Human Male BAC"
1.56
/misc_feature
/notes="assembly_fragment"
clone_end:T7
vector_side:right"
157.1191
/misc_feature
/notes="assembly_fragment"
1292.2432
/misc_feature
/notes="assembly_fragment"
2533.3742
/misc_feature
/notes="assembly_fragment"
3843.4921
/misc_feature
/notes="assembly_fragment"
5022.6073
/notes="assembly_fragment"
clone_end:SP6
vector_side:left"
6174.7552
/misc_feature
/notes="assembly_fragment"
7653.9001
/misc_feature
/notes="assembly_fragment"
9102.10622
/misc_feature
/notes="assembly_fragment"
10723.11836
/misc_feature
/notes="assembly_fragment"
11937.13398
/misc_feature
/notes="assembly_fragment"
13499.14949
/misc_feature
/notes="assembly_fragment"
15050.16260
/misc_feature
/notes="assembly_fragment"
16361.18265
/notes="assembly_fragment"

Query Match 94.2%; Score 672.8; DB 2; Length 175046;
Best Local Similarity 98.2%; Pred.No. 4.9e-89;
Matches 699; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 2 GCCAGAAAAAGTATTATTTTCTATTAACATCTTCTCAAGCATTTATTCCT 61
Db 43357 GCCAGAAAAAGTATTATTTTCTATTAACATCTTCTCAAGCATTTATTCCT 43298

QY 62 ATATCTCACTGAATTTAAGAAATACATTAGTATTAGAAAACTAGGAAAAAGATPAA 121
Db 43297 ATATCTCACTGAATTTAAGAAATACATTAGTATTAGAAAACTAGGAAAAAGATPAA 43238

QY 122 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATTAACAAGGACTGAAACGTTATA 181
Db 43237 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATTAACAAGGACTGAAACGTTATA 43178

QY 182 AATTGAATGAGATTATAATTGAAAACTGCATCTGAAGCAAACTTTATGTTCAATTA 241
Db 43177 AATTGAATGAGATTATAATTGAAAACTGCATCTGAAGCAAACTTTATGTTCAATTA 43118

QY 242 TNCTTAATGATGGTGTATTATGACTAATACACTGATTTTCAAGAGGAAACCACTGTTA 301
Db 43117 TTCTTAATGATGGTGTATTATGACTAATACACTGATTTTCAATTAAGGAAACCACTGTTA 43058

QY 302 AAAATATTTTATTATTAATAAAGCCCTGTGTTCAAGCTCTGATCATATTCTTTTATT 361
Db 43057 AAAATATTTTATTATTAATAAAGCCCTGTGTTCAAGCTCTGATCATATTCTTTTATT 42998

QY 362 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 421
Db 42998 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 421
```

```
Db 42997 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 42938
QY 422 TCTCACTAATTTTAAANAACCTATTGAGAAATGATTAATGACATGAAGTGCAACAATA 481
Db 42937 TCTCACTAATTTTAAAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAATA 42878
QY 482 TTTACTGGCCAGCTGTGGCATTTGTCTTACTTAGTTCTCCCAAGGAAAACTCTTAA 541
Db 42877 TTTACTGGCCAGCTGTGGCATTTGTCTTACTTAGTTCTCCCAAGGAAAACTCTTAA 42819
QY 542 ATTGAATCTTCAGACGAATATCCTTAATATTAATCTTTGTAAGCAAAACAAAGCTTTT 601
Db 42818 ACTGAATCTTCAGACGAATATCCTTAATATTAATCTTTGTAAGCAAAACAAAGCTTTT 42759
QY 602 GTTACATAGTTCTTTGGGATTCTTCTCTAATTTTATTCGAAACTCAATTTTACC 661
Db 42758 GTTACATAGTTCTTT-GGATTTTACTGTTCTTAATTTTATTCGAAACTCAATTTTACC 42700
QY 662 CCAGACCATTAATACCATATTAACTTTGTTNGACAGAGTTGTTGCCAATTC 713
Db 42698 CCAGACCATTAATACCATATTAACTTTGTTNGACAGAGTTGTTGCCAATTC 42648

RESULT 7
AL445183/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-334A14 on chromosome 1, complete
sequence.
ACCESSION
AL445183
VERSION
AL445183.19 GI:20068427
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1
AUTHORS
Wallis,J.
TITLE
Direct Submision
JOURNAL
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:17939714.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP11-334A14 is from the library RPC1-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.
FEATURES
Source
1.193774
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

```

/chromosome="1"
/clone="RP11-334A14"
/clone_lib="RPC1-11.2"
BASE COUNT      50815 a 45351 c 44031 g 53577 t
ORIGIN

```

Query Match	94.2%	Score 672.8;	DB 9;	Length 193774;
Best Local Similarity	98.2%;	Pred. No. 4.7e-89;		
Matches 699; Conservative	0;	Mismatches 11;	Indels 2;	Gaps 2;

QY	2	GCCAGAAAAAGTATTTTAAATTTTCTAATTAACATTTCTTCTCAAGCATTTATTTATCCT	61
Db	98224	GCCAGAAAAAGTATTTTAAATTTTCTAATTAACATTTCTTCTCAAGCATTTATTTATCCT	98165
QY	62	ATATCTCACTGAATTTTAAAGAAATTAACATTAGTATTAGAAAAAACTAGAAAAAAGATATA	121
Db	98164	ATATCTCACTGAATTTTAAAGAAATTAACATTAGTATTAGAAAAAACTAGAAAAAAGATATA	98105

Qy	122	TCACAGATAATTAACTTACATGAAAAAGAAAAATTATTAACAAAGCACTGAGAACGTTATA	181
Db	98104	TGCAGATAATTAACTTACATGAAAAAGAAAAATTATTAACAAAGCACTGAGAACGTTATA	98045

Oy	182	AATTGAATGAGATTATAATTGAAAACTCGATCTGAAAGCAACTTTATTGTCCAATTA	241
D6	98044	AATTGAATGAGATTATAATTGAAAACTCGATCTGAAAGCAACTTTATTGTCCAATTA	97985

[illegible]

QY	302	AAAAATATTTTATTTTAAAAATAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTT	361
Db	97924	AAAAATATTTTATTTTAAAAATAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTT	97865

Oy	362	TGATTTGGGAANAATACTGTTCGTAGTGCATGAAATGCAAAATTTT	TAGATTTTTTA	421
Db	97864	TGATTTGGGAAGAATACTGTTCGTAGTGCATGAAATGCAAAATTTT	TAGATTTTTTA	97805

OY		422	TCTCACTAA TTTTAANNACTATTGAGAA TTGATTAATGACATGAAGTGACAACA CTTAA	481
Dd		97804	TCTCACTAATTTTAA GACTATTGAGAA ATTGATTAA TGATCATA GTGAAGTGACA CAACACTTAA	97745

Oy		482 TTACTGGCCAGCTGTGTCATTGTTTCTTA CTTAGTTCGCCAAGGA AACTTTAA	541
Dδ		97744 TTACTGGCCAGCTGTGTCATTGTTTCTTA CTTAGTTCGCCA -GGAAAC TTTAA	97686

Qy	542	ATTGAATCTTCAGCAGCAATAATCTTAAATATACTTTGTAGCAAAA	CAAAAGCTTTTT	601
Db	97685	ACTGAATCTTCAGCAGCAATAATCTTAAATATACTTTGTAGCAAAA <td>CAAAAGCTTTTT</td> <td>97626</td>	CAAAAGCTTTTT	97626

Qy 602 GTTTACATAGTCTCTTGGGATTTTACTGTCTCCTAATTTTAACTGAAACTCAATTTTACC 661
|||||
Dd 97625 GTTTACATAGTCTCTT -GGATTTTACTGTCTCCTAATTTTAACTGAAACTCAATTTTACC 97567

Qy	662	CCAGACATAATTACCATATTAACTTGTNTGCACAGTTGTTGCCAATTC	713
Db	97566	CCAGACATAATTACCATATTAACTTGTATAATGCACAGTTGATGCAATTC	97515

RESULT 8
BC005911/c

DEFINITION	Homo sapiens, sterol carrier protein 2, clone MGC:14505
IMAGE	IMAGE:4287946, mRNA, complete cds.
ACCESSION	BC005911

KEYWORDS	ORGANISM
MG.	Homo sapiens (human)
SOURCE	Homo sapiens

REFERENCE
AUTHORS
1 (bases 1 to 1439)
Strausberg, R.

TITLE	Direct Submission
JOURNAL	Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

Direct Submission
Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaphs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 21 Row: n Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 432978.

FEATURES	Location/Qualifiers
source	1. .1439

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocustID:6342"
/db_xref="taxon:9606"
/clone="MGC:14505 IMAGE:4287946"
/tissue_type="Brain, primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
74. .505
DSDS

```

```

/codon_start=1
/product="sterol carrier protein 2"
/protein_id="AAH05911.1"
/db_xref="GI:13543503"
/translation="MGFPEAASSFRTHOIEAVPTSSASDGFKANLVKEIEKKL.EEEG
EQFVKIIGIEFAFKYKDFGEGKEATWVDVKNKGSVLPNSDKKADCTITMADSDFLA
LMTGKNPQSAFEOGKLIKTNMGLAMKLNLOLQGNNAKL"

```

Query Match	93.5%;	Score 667.6;	DB 9;	Length 1439;
Best Local Similarity	97.9%;	Pred. No. 1.5e-87;		
Matches 695;	Conservative	0;	Mismatches 13;	Indels 2;
				Gaps 2;

Qy	4	CAGAAAAAGTATTTTAAATTTCTATTAACATTCTCTCAAGCATTATTTATCTAT	63
Db	1409	CGGAAAAAGTTATTTTAAATTTCTATTAACATTCTCTCAAGCATTATTTATCTAT	1356

Qy	64	ATCTCACTGAATTTTAA	GAAATAACATTAGTATTAGAAAACTAGAAAAAAGATAAATG	123
Db	1349	ATCTCACTGAATTTTAA <td>GAAATAACATTAGTATTAGAAAACTAGAAAAAAGATAAATG</td> <td>1290</td>	GAAATAACATTAGTATTAGAAAACTAGAAAAAAGATAAATG	1290

Qy	124	CAGATAATTAACTTACATGAAAAAGAAAAATTATCAAGAAGCTGAAACGTTATAAA	183
Db	1289	CAGATAATTAACTTACATGAAAAAGAAAAATTATCAAGAAGCTGAAACGTTATAAA	1230

QY	184	TTGAAATGAGATTATTAATTTGAAACTGCATCTGAAAGCAACTTATTTGTTCAATTATN	243
Db	1229	TTGAAATGAGATTATTAATTTGAAACTGCATCTGAAAGCAACTTATTTGTTCAATTATT	1170

244 CTTAATGATGTTTATGACTAATACACTGATTTTTCAGAAGGAACCCATGTTAA 303

QY 304 AATATTTTATTAAAAATAAGCTGTGTTCAAGCTCTGATCATATTTCTTTAATTTTG 363
|||
Db 1109 AATATTTTATTAAAAATAAGCTGTGTTCAAGCTCTGATCATATTTCTTTAATTTTG 1050

Best Local Similarity 97.1%; Pred. No. 2.3e-79; Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

QY	15	ATTTTAATTTCTATTAACATCTTCTCAAGCATTTATTTATCCTATATCTCAGTGA	74
Db	1	ATTTTAATTTCTATTAACATCTTCTCAAGCATTTATTTATCCTATATCTCAGTGA	60
QY	75	TTTTAGAATAACATAGTATTTAGAAAACTAGGAAAAAGATAAATGCAGATAATTAA	134
Db	61	TTTTAANAATAACATTAGTATTTAGAAAACTAGGAAAAAGATAAATGCAGATAATTAA	120
QY	135	ACTTACATGAAAAAGAAAAATTATTAACAAGGAGCTGAGAACGTTATAATTGAAATGAGA	194
Db	121	ACTTACATGAAAAAGAAAAATTATTAACAAGGAGCTGAGAACGTTATAATTGAAATGAGA	180
QY	195	TTATTAATTTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG	254
Db	181	TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG	240
QY	255	TGTTTATGACTAATACATCGATTTTTCAGAAGGAAACCATGTTAAAAATATTTTAT	314
Db	241	TGTTTATGACTAATACATCGATTTTTCAGAAGGAAACCATGTTAAAAATATTTTAT	300
QY	315	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTGATTGGGAANA	374
Db	301	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTGATTGGGAANA	360
QY	375	AAATACGTGTTCTGATAGCATGAAATGCCAAATTTTGAATTTTAACTCTCACTAATTTT	434
Db	361	AAATACGTGTTCTGATAGCATGAAATGCCAAATTTTGAATTTTAACTCTCACTAATTTT	420
QY	435	AANAACATTTGAGAAATTTGATTAATGACATGAAGTGACACACTAATTACTGGCCAGCT	494
Db	421	AAGAACTATTGAGAAATTTGATTAATGACATGAAGTGACACACTAATTACTGGCCAGCT	480
QY	495	GTTGGCATTTGTTTCTTACTTAGTTCTCCCAAGGAAACCTTAAATTGAATCTTCAG	554
Db	481	GTTGGCATTTGTTTCTTACTTAGTTCTCCCAAGGAAACCTTAAATTGAATCTTCAG	539
QY	555	CAGAAATAATCCTTAATAATACCTTTGTAAGCAAAAACAAAGCTTTTGTGTACATAGTTC	614
Db	540	CNGAATAA-CCTTAATAATACCTTTGTAAGCAAAAACAAAGCTTTTGTGTACATAGTTC	597
QY	615	TTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACCCAGACCATTAAT	674
Db	598	TTT-GGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACCCAGACCATTAAT	656
QY	675	ACCATA-TTAACTTTGTTNTGCACAGTTGTT	704
Db	657	ACCCTATTTAACCTTTGTTNTGCACAGTTGTT	687

RESULT 11
S52450/c 1229 bp mRNA linear PRI 08-MAY-1993
LOCUS sterol carrier protein 2 [human, liver, mRNA, 1229 nt].
DEFINITION S52450
ACCESSION S52450.1 GI:263550
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Yamamoto, R.
TITLE Localization of human sterol carrier protein 2 gene and cDNA
expression in COS-7 cell
JOURNAL Hokkaido Igaku Zasshi 67 (6), 839-848 (1992)
MEDLINE 93131254
PUBMED 1483685
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbon 122504] from the original journal article.
This sequence comes from Fig. 1.

Map location: 1.
FEATURES location/Qualifiers
source 1..1229
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1229

gene /gene="sterol carrier protein 2, SCP2"

CDS

84..515 /gene="sterol carrier protein 2, SCP2"

/note="This sequence comes from Fig. 1; SCP2"

/product="sterol carrier protein 2"

/protein_id="AAB24921.1"

/db_xref="GI:263551"

/translation="MGFPEAASSFRTHQIEAVPTSSASDGFKANIVKEIEKLEBEG
EQFVKKIGIFAFKVKDGPGEKEDTWVDVKNQGSVLPNSDKADCTITMAASDFLA
LMTGKMPQSAFFQGPLKITGNMGLAMKLNQLQPGNAKL"

BASE COUNT 409 a 200 c 255 g 365 t
ORIGIN

Query Match 66.5%; Score 474.6; DB 9; Length 1229;
Best Local Similarity 97.1%; Pred. No. 1.2e-59;
Matches 502; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY	195	TTATTAATTTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG	254
Db	1228	TTTTTTTGAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG	1169
QY	255	TGTTTATGACTTAATACTGATTTTTCAGAAGAAAAACCATGTTAAAAATATTTTAT	314
Db	1168	TGTTTATGACTTAATACTGATTTTTCAGAAGAAAAACCATGTTAAAAATATTTTAT	1109
QY	315	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTGATTGGGAANA	374
Db	1108	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTGATTGGGAANA	1049
QY	375	AAATACGTGTTCTGATAGCATGAATGCAAAATTTTGAATTTTAACTCTCACTAATTTT	434
Db	1048	AAATACGTGTTCTGATAGCATGAATGCAAAATTTTGAATTTTAACTCTCACTAATTTT	989
QY	435	AANAACATTTGAGAAATTTGATTAATGACATGAAGTGACACACTAATTACTGGCCAGCT	494
Db	988	AAGAACTATTGAGAAATTTGATTAATGACATGAAGTGACACACTAATTACTGGCCAGCT	929
QY	495	GTTGGCATTTGTTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAATTTGAATCTTCAG	554
Db	928	GTTGGCATTTGTTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAATTTGAATCTTCAG	870
QY	555	CAGAAATAATCCTTAATAATACCTTTGTAAGCAAAAACAAAGCTTTTGTGTACATAGTTC	614
Db	869	CAGAAATAATCCTTAATAATACCTTTGTAAGCAAAAACAAAGCTTTTGTGTACATAGTTC	810
QY	615	TTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACCCAGACCATTAAT	674
Db	809	TTT-GGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACCCAGACCATTAAT	751
QY	675	ACCATAATTAACCTTTGTTNTGCACAGTTGTTGCCAAT	711
Db	750	ACCATAATTAACCTTTGTTNTGCACAGTTGTTGCCAAT	714

RESULT 12
HUMSTEA/c 1219 bp mRNA linear PRI 06-DEC-1993
LOCUS Human sterol carrier protein-2 (SCP-2) mRNA, complete cds.
DEFINITION M55421
ACCESSION M55421.1 GI:432978
VERSION
KEYWORDS sterol carrier protein-2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1219)
AUTHORS Yamamoto,R., Kallen,C.B., Babalola,G.O., Rennert,H.,
TITLE Billheimer,J.T. and Straus,J.F. III.
Cloning and expression of a cDNA encoding human sterol carrier
protein 2
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (2), 463-467 (1991)
MEDLINE 9110550
PUBMED 1703300
REFERENCE 2
AUTHORS Vesa,J., Hellsten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T.,
TITLE Mead,S., Cowell,J.K., Straus,J.F.III. and Peltonen,L.
Assignment of sterol carrier protein X/sterol carrier protein 2 to
1p32 and exclusion as the causative gene for infantile neuronal
ceroid lipofusiosis
JOURNAL Unpublished
COMMENT On Dec 6, 1993 this sequence version replaced gi:432972.
Original source text: Human female liver.
FEATURES
Source location/Qualifiers
1..1219
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="1p32"
/tissue_type="liver"
1..1219
/gene="SCP-2"
84..515
/gene="SCP-2"
/codon_start=1
/product="sterol carrier protein-2"
/protein_id="AA03559.1"
/db_xref="GI:432973"
/translation="MGFPEAASSFRTHQIEAVPTSSASDGFKANLVFKEIEKLEEG
EQFVKIIGIFAFKVKDGPGEKATWVDVKNKGSVLPNSDKADCTTMADSDFLA
LMTGKNPQSAFPGKLIKITGMGLAMKLNQLQPGNAKL"
84..143
/gene="SCP-2"
144..512
/gene="SCP-2"
/product="unnamed"
782..787
/gene="SCP-2"
/evidence=not_experimental
1062..1067
/gene="SCP-2"
/evidence=not_experimental
1107..1112
/gene="SCP-2"
/evidence=not_experimental
1190..1195
/gene="SCP-2"
/evidence=experimental
1219
/gene="SCP-2"
/evidence=experimental
BASE COUNT 402 a 195 c 255 g 367 t
ORIGIN
Query Match 64.4%; Score 459.8; DB 9; Length 1219;
Best Local Similarity 97.3%; Pred. No. 1.7e-57;
Matches 497; Conservative 0; Mismatches 11; Indels 3; Gaps 3;
QY 203 TGAAGCTGCATCTGAAAGCAACTTATTGTTCAATTATNCTTAATGATGGTGTAT 262
Db 1219 TGAAGCTGCATCTGAAAGCAACTTATTGTTCAATTATNCTTAATGATGGTGTAT 1160
QY 263 GACTAATACACTGATTTTCAAGAGAAACCATGTTAAATAATTTTATTTTAAAA 322
Db 1159 GACTAATACACTGATTTTCAAGAGAAACCATGTTAAATAATTTTATTTAAAA 1100
QY 323 TAAGCTGTGTCAAGCTCTGATCATATTCTTTTATTTGATTTGGAAANAATACTG 382
Db 1099 TAAGCTGTGTCAAGCTCTGATCATATTCTTTTATTTGATTTGGAAANAATACTG 1040

QY 383 TTTCTGATAGCATGAATGCCAAATTTTGTAGATTTTAACTCTACTAATTTAANAAC 442
Db 1039 TTTCTGATAGCATGAATGCCAAATTTTGTAGATTTT-ATCTACTAATTTAAGAACTA 981
QY 443 TTGAGAAATTGATTAAAGACATGAAGTGCACAACTAATTACTGGCCAGCTGTGGCAT 502
Db 980 TTGAGAAATTGATTAAAGACATGAAGTGCACAACTAATTACTGGCCAGCTGTGGCAT 921
QY 503 TGTGTTTCTTACTAGTCTCTCCAGGAGAAACCTTAAATTGAATCTTCAGAGAAATA 562
Db 920 TGTGTTTCTTACTAGTCTCTCCAGGAGAAACCTTAAATTGAATCTTCAGAGAAATA 862
QY 563 TCCTTAATATACTTTGTAAGCAAAACAAAGCTTTTGTGTACATAGTTCTTGGGAT 622
Db 861 TCCTTAATATACTTTGTAAGCAAAACAAAGCTTTTGTGTACATAGTTCTT-GGAT 803
QY 623 TTTACTGTCTCTAATTTTATTTCTGAACCTCAATTTTACCCAGACCATTAATACCATAT 682
Db 802 TTTACTGTCTCTAATTTTATTTCTGAACCTCAATTTTACCCAGACCATTAATACCATAT 743
QY 683 AACTTGTGTGACAGTGTGTTGCCAATTC 713
Db 742 AACTTGTGTAATGCACAGTTGTATGCAATTC 712
RESULT 13
ARI46582
LOCUS ARI46582 843 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 22 from patent US 6218521.
ACCESSION ARI46582
VERSION ARI46582.1 GI:15109771
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 843)
AUTHORS Obata,Y.
TITLE Isolated nucleic acid molecules associated with gastric cancer and
methods for diagnosing and treating gastric cancer
JOURNAL Patent: US 6218521-A 22 17-Apr-2001;
FEATURES location/Qualifiers
Source 1..843
/organism="unknown"
BASE COUNT 237 a 184 c 89 g 215 t 118 others
ORIGIN
Query Match 61.1%; Score 436.2; DB 6; Length 843;
Best Local Similarity 73.1%; Pred. No. 5e-54;
Matches 516; Conservative 0; Mismatches 187; Indels 3; Gaps 3;
QY 2 GCCAGAAAAAGTTATTTAATTTCTATTAACAATTTCTCAAGCATTTATTTATCCT 61
Db 2 GCCAAAAAANTTATTNAATTTCTATTAANCNTCTCCNCAANCAATTATTNACCT 61
QY 62 ATATCTCACTGAATTTTAAGAAATAACATTAGATTAGAAAACTAGAAAAAAGATAAA 121
Db 62 ATNNCNCNCNGANTTTNANAANAANTACCTTTNNNTTTAAAAAACTNGAAAAAATAAT 121
QY 122 TGCAGATAATTAAACTTACATGAAGAAAGAAATTAACAAGAGACTGAGAACGTATA 181
Db 122 NGCAATANTTAACCTTCTGAAAGAAATTTNACCAANGAGACNGAAANCNTNTA 181
QY 182 AATTGAATGAGATTATTAATTGAAAACATGCATCTGAAAGCAAACCTTATTTGTTCAATTA 241
Db 182 ATTNGAANTNAATTATANTTNGAAGAACGCGCNCNGAAACCAANCTNATGTCCAATTA 241
QY 242 TNCCTAATGATGTGTTTATGACTAATACACTGATTTTCAAGAGAAACCCATGTTA 301
Db 242 TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAAATANGGAANCCNNNTTA 301
QY 302 AAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 361
Db 302 AAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 361

Db 302 AAANTNTTTNATTTTAAATAACCCNGTNTCCAACCCCGATCANATTCTTMTATT 361
QY 362 TGATTGGGAANAATACTGTTCTGATGACATGAATGCAAAATTTTAGATTTTAA 421
Db 362 GGATTGGGGAATAAATNCNGTTCNNATACCNNGAANNCAAANTTTTAAATTTTAA 421
QY 422 TCTCACTAATTTTAAANAATACTATTGAGAAATTGATTATGACATGAGTGACACACTAA 481
Db 422 CCCCCCTANTTTTAAANCTATNGAAANTNGATTANNGACTGAAATTGC-CAACCCCTAN 480
QY 482 TTACTGGCCAGCTGTTGGCATGTGTTCTTACTAGTTCTCCCAAGGAAAACTCTTAA 541
Db 481 TTNCNGGCCACNGTGGCNTNGTNTCTTACTTANTCCCCCAA-GGAAANNCCTTAA 539
QY 542 ATTGAATCTTCAGCAGATAATCCTTAATATACTTTGTAAGCAAAACAAGCTTTT 601
Db 540 NCNGAANCTCCNCAATAACCTTANATATCCTTGTAACCAAAACCAACCTTTTIN 599
QY 602 GTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACC 661
Db 600 GTTACNTA-NTCCTTGGGATTTAAGGGTCCCAATTNATCNGAACCCANTTTTCCC 658
QY 662 CCAGACCATTAATTAACCATTAATTTGTTNGCACAGTTGTTGC 707
Db 659 CCNAACCATANTTACCATTTTACTTGTGTAAGCNCAGTNGTTGC 704

RESULT 14

BD079831 843 bp DNA linear PAT 27-AUG-2002
LOCUS Cancer-associated nucleic acids and polypeptides.
DEFINITION BD079831
ACCESSION BD079831
VERSION BD079831.1 GI:22625434
KEYWORDS JP 2001516009-A/497.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Old.L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I.,
Ochare,M., Obata,Y., Pfeundschnuh,M., Tureci,O. and Sahin,U.
Cancer-associated nucleic acids and polypeptides
JOURNAL Patent: JP 2001516009-A 497 25-SEP-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH

COMMENT OS Homo sapiens (human)
PN JP 2001516009-A/497
PD 25-SEP-2001
PF 15-JUL-1998 JP 2000503425
PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR
10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR
11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD
J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG
CHEN, IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFEUNDSCNUH, PI
OZLEM TURECI,
PI UGUR SAHIN

PC GOIN33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC
A61P35/00,
PC C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
PC C12N15/00
CC Cancer-associated nucleic acids and polypeptides. FH Key
LOCATION/Qualifiers
FT source 1..843
FT /Organism='Homo sapiens (human)'.
FEATURES
source 1..843
Location/Qualifiers
1..843
/Organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 237 a 184 c 89 g 215 t 118 others
ORIGIN

Query Match 61.1%; Score 436.2; DB 6; Length 843;
Best Local Similarity 73.1%; Pred. No. 5e-54;
Matches 516; Conservative 0; Mismatches 187; Indels 3; Gaps 3;

QY 2 GCCAGAAAAAGTTATTTTAAATTTCTATTAAACATCTCTCAAGCATTATTATCTT 61
Db 2 GCCAAAAAANTTATTNAAATTTCTATTAANCNTCCCTCNCAANCAATTATTNACCT 61
QY 62 ATATCTACTGAATTTTAAAGAAATAACATTAGATTAGAAAACTAGAAAAAGATAAA 121
Db 62 ATNNCNCNCGANTTTNANAANAATACCTTNTNTTAAAAAACCCTNGAAAAAATAAT 121
QY 122 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAAGAGACTGAGAACGTATA 181
Db 122 NGCAAAATANTTAACCTTNCCTTGAAAAANGAAAAATTNTTACCAANGAACNGAAANCNTNTA 181
QY 182 AATTGAAATGAGATTATAATTGAAAACTGCATCTGAAAAAGCAAACTTATTGTTCAATTA 241
Db 182 ATTNGAANTNAAATTATANTNGAAAAANGCGCNCNNGAAAAACAANCTTATGTTCAATTA 241
QY 242 TNCCTAATGATGTTGTTTATGACTAATACATGATGATTTTCAAGAGAAAGCAATGTTA 301
Db 242 TCTTNAANGAGGNTTTNANNACTAATNCCNGAATTTTCCATANGAANGCAANCCNNNTTA 301
QY 302 AAAATATTTTATTTTAAATAAAGCCCTGTGTTCAAGCTCTGATCATATTCTTTAATT 361
Db 302 AAANTNTTTNATTTTAAATAAACCNGTNTCCAACCCCGATCANATTCTTTNATTT 361
QY 362 TGATTGGGAANAATACTGTTCTGATGACATGAATGCAAAATTTTAGATTTTAA 421
Db 362 GGATTGGGGAATAAATNCNGTTCNNATACCNNGAANNCAAANTTTTAAATTTTAA 421
QY 422 TCTCACTAATTTTAAANAATACTATTGAGAAATTGATTATGACATGAGTGACACACTAA 481
Db 422 CCCCCCTANTTTTAAANCTATNGAAANTNGATTANNGACTGAAATTGC-CAACCCCTAN 480
QY 482 TTACTGGCCAGCTGTTGGCATGTGTTCTTACTAGTTCTCCCAAGGAAAACTCTTAA 541
Db 481 TTNCNGGCCACNGTGGCNTNGTNTCTTACTTANTCCCCCAA-GGAAANNCCTTAA 539
QY 542 ATTGAATCTTCAGCAGATAATCCTTAATATACTTTGTAAGCAAAACAAGCTTTT 601
Db 540 NCNGAANCTCCNCAATAAACCCTTANATATCCTTGTAACCAANCAAAACCTTTTIN 599
QY 602 GTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACC 661
Db 600 GTTACNTA-NTCCTTGGGATTTAAGGGTCCCAATTNATCNGAACCCANTTTTCCC 658
QY 662 CCAGACCATTAATTAACCATTAATTTGTTNGCACAGTTGTTGC 707
Db 659 CCNAACCATANTTACCATTTTACTTGTGTAAGCNCAGTNGTTGC 704

RESULT 15

HUMSCP2B/c 1500 bp mRNA linear PRI 06-DEC-1993
LOCUS Human sterol carrier protein, 2 mRNA, complete cds.
DEFINITION HUMSCP2B
ACCESSION M75884
VERSION M75884.1 GI:432976
KEYWORDS sterol carrier protein-2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 1500)
He,Z., Yamamoto,R., Furth,E.E., Schantz,L.J., Naylor,S.L.,
George,H., Billheimer,J.T. and Straus,J.F. III.
cDNAs encoding members of a family of proteins related to human
sterol carrier protein 2 and assignment of the gene to human
chromosome 1 p21----pter
JOURNAL DNA Cell Biol. 10 (8), 559-569 (1991)
MEDLINE 92029618
PUBMED 1718316

REFERENCE	2 (bases 1 to 1500)
AUTHORS	Vessa,J., Hellsten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T., Mead,S., Cowell,J.K., Strauss,J.F.III, and Peltonen,L.
TITLE	Assignment of sterol carrier protein X/sterol carrier protein 2 to 1p32 and exclusion as the causative gene for infantile neuronal ceroid lipofusiosis
JOURNAL	Unpublished
COMMENT	On Dec 6, 1993 this sequence version replaced gi:337996. Original source text: Human liver cDNA to mRNA.
FEATURES	location/Qualifiers
source	1..1500
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/map="1p32"
	/cissue_type="liver"
	/clone_1ib="Clontech HL-1001b;HL-1115b"
gene	1..1500
	/gene="SCP-2"
CDS	82..951
	/gene="SCP-2"
	/codon_start=1
	/product="sterol carrier protein-2"
	/protein_id="AA03558.1"
	/db_xref="GI:432977"
	/translation="MTDLPSSFEEKSIKRVGFDMSKEARKCYEKSGLTPNDIDVI ELHDCFSTNELLYEALGLCPBGQATLVDRDNTYGGKVINPSSGLISKHPLGAT GLAOCAELCWOLRGEAGKRVGAKVALOHNLGGAVVTLVKMGFPEAASSERTHO IEAVPTSSASDGFKANLVFKEIEKLEEEGEQFVKIGIIPAFKVKDGPGEKATWV DVKNGKSVLPNSDKKADCTITMADSDFLALMTGKMPQSAFQGLKITGNMGLAMK LQNLQLQPGNAKL"
	1218..1223
	/gene="SCP-2"
	/note="no polyA or polyA signal was found at the end of this cDNA clone, indicating that the mRNA sequence may be partial at the 3' end"
	/evidence=not experimental
BASE COUNT	482 a 254 c 342 g 422 t
ORIGIN	
Query Match	44.6%; Score 318.4; DB 9; Length 1500;
Best Local Similarity	96.9%; Pred. No. 4.3e-37;
Matches	344; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
QY	359 TTTTGATTGGGAANAATACTGTTCTGATAGCATGAAATGCAAAATTTTAGATTTT 418
Db	1500 TTTTGATTGGGAAGAAATACTGTTCTGATAGCATGAAATGCAAAATTTTAGATTTT 1441
QY	419 TAACTCTCACTAATTTTAAANAATACTGAGAAATTGATTAATGACATGAAGTCACAC 478
Db	1440 TAACTCTCACTAATTTTAAAGAACTATGAGAAATTGATTAATGACATGAAGTCACAC 1381
QY	479 TAACTACTGGCCAGCTGTGGCATTTGTTCTTACTTAAGTTCTCCCAAGGAAACTCT 538
Db	1380 TAACTACTGGCCAGCTGTGGCATTTGTTCTTACTTAAGTTCTCCCAAGGAAACTCT 1322
QY	539 TAAATTGAATCTTCAGCAGATAATCCTTAATATATCTTTGTAAGCAAAACAAAGCTTT 598
Db	1321 TAAACTGAATCTTCAGCAGATAATCCTTAATATATCTTTGTAAGCAAAACAAAGCTTT 1262
QY	599 TTGTTTACATAGTTCCTTTGGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTT 658
Db	1261 TTGTTTACATAGTTCCTT-GGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTT 1203
QY	659 ACCCCAGACCATAATTACCATATTAACCTTGTGTTNGACAGTGTGGCAATTC 713
Db	1202 ACCCCAGACCATAATTACCATATTAACCTTGTGTTNGACAGTGTGATGCAATTC 1148

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:55:53 ; Search time 1924.36 Seconds
(without alignments)
9017.749 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 714
Sequence: 1 cgcacagaaaaagtcattctta.....cacagctgttcgcaattca 714

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	94.3	769	12 BQ014192	BQ014192 UI-H-ED1-
2	672.8	94.2	761	12 BM997078	BM997078 UI-H-ED0-
3	671.8	94.1	821	9 AI640146	AI640146 wa29f11.x
4	670.4	93.9	770	9 AI826287	AI826287 wk33f07.x

5	669.6	93.8	756	12 BM985376	BM985376 UI-CF-EC1
6	668.4	93.6	738	12 BQ045161	BQ045161 UI-CF-EN1
7	667	93.4	786	9 AW052045	AW052045 wx25e05.x
8	664.4	93.1	751	12 BQ008197	BQ008197 UI-H-ED1-
9	664.4	93.1	759	13 BU619112	BU619112 UI-H-FH1-
10	664.2	93.0	721	12 BM968746	BM968746 UI-CF-DU1
11	663.2	92.9	734	14 CA424156	CA424156 UI-H-FE1-
12	661.4	92.6	762	13 BU627147	BU627147 UI-H-FG0-
13	660.6	92.5	734	13 BU933572	BU933572 AGENCOURT
14	653.4	91.5	712	13 BU619082	BU619082 UI-H-FH1-
15	651.2	91.2	750	12 BQ009270	BQ009270 UI-H-ED1-
16	649	90.9	905	13 BX414370	BX414370 BX414370
17	648.8	90.9	901	13 BX392565	BX392565 BX392565
18	645.2	90.4	840	9 AA659242	AA659242 nu12h08.8
19	645	90.3	953	9 AL578584	AL578584 AL578584
20	644.4	90.3	837	9 AI566109	AI566109 tns3e08.x
21	638.6	89.4	714	9 AI675901	AI675901 wc06h04.x
22	636.8	89.2	822	9 AV729461	AV729461 AV729461
23	634.4	88.9	793	9 AI565988	AI565988 tns2a05.x
24	631.2	88.4	672	12 BQ016778	BQ016778 UI-H-D10-
25	626.8	87.8	784	12 BI861076	BI861076 603390807
26	623.2	87.3	876	9 AI831751	AI831751 wj40a10.x
27	622.2	87.1	699	10 BE739622	BE739622 601556314
28	620.6	86.9	824	10 BE738457	BE738457 601572489
29	619.6	86.8	641	9 AW173415	AW173415 xj06d03.x
30	619.6	86.8	728	9 AI831053	AI831053 wj62c05.x
31	615.4	86.2	749	9 AI765940	AI765940 wh69d12.x
32	612.8	85.8	671	9 AI683310	AI683310 tw50e02.x
33	609.6	85.4	1032	9 AL551280	AL551280 AL551280
34	607.2	85.0	648	12 BM311898	BM311898 1g63e08.y
35	605.4	84.8	673	9 AW512844	AW512844 xm04g12.x
36	604.2	84.6	820	10 BF966918	BF966918 602286426
37	603	84.5	637	13 BQ549593	BQ549593 1k88g10.x
38	593.8	83.2	763	9 AI889069	AI889069 cd17a12.x
39	592	82.9	657	9 AI627276	AI627276 cy74e03.x
40	590.6	82.7	646	9 AW518060	AW518060 xx80e07.x
41	587.8	82.3	623	12 BM511625	BM511625 1j51a05.x
42	587	82.2	627	9 AI224925	AI224925 qi31c04.x
43	585.6	82.0	616	9 AW611897	AW611897 hg90b03.x
44	584.8	81.9	744	10 BG569206	BG569206 602588480
45	582.6	81.6	702	9 AI431843	AI431843 ci25h04.x

ALIGNMENTS

RESULT 1 BQ014192 769 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-ED1-axs-g-24-0-UI.81 NCI_CGAP_ED1 Homo sapiens CDNA clone
DEFINITION IMAGE:5833007 3', mRNA sequence.

ACCESSION BQ014192
VERSION BQ014192.1 GI:19739093
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 769)

JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this CDNA sequence: 317-343, >AT_richlow_complexity (matched complement)

Seq primer: M13 FORWARD
POLYA=Yes.
location/Qualifiers
1. 769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5833007"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI_CGAP_ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_LIB=UI-H-ED1
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"
BASE COUNT 272 a 110 c 100 g 285 t 2 others
ORIGIN
Query Match 94.3%; Score 673; DB 12; Length 769;
Best Local Similarity 98.3%; Pred. No. 7e-70;
Matches 698; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
QY 2 GCCAGAAAAAGTTATTTTAAATTTCTATTAACATCTCTCAAGCATTTTATCCCT 61
DB 20 GCCAGAAAAAGTTATTTTAAATTTCTATTAACATCTCTCAAGCATTTTATCCCT 79
QY 62 ATATCTCACTGAATTTTAAAGAAATAACATGATTTAGAAAAACTAGAAAAAGATAAA 121
DB 80 ATATCTCACTGAATTTTAAAGAAATAACATGATTTAGAAAAACTAGAAAAAGATAAA 139
QY 122 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATA 181
DB 140 TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATA 199
QY 182 AATTGAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACCTTAATTGTTCAATT 241
DB 200 AATTGAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACCTTAATTGTTCAATT 259
QY 242 TNCCTTAATGATGGTGTCTTATGACTAATACATGATTTTCAAGAGGAAACCCTGTTA 301
DB 260 TTCTTAATGATGGTGTCTTATGACTAATACATGATTTTCAATAGGAAACCCTGTTA 319
QY 302 AAAATATTTTATTTTAAATAAAGCCGTGTCAAGCTCTGATCATATTTCTTTATTT 361
DB 320 AAAATATTTTATTTTAAATAAAGCCGTGTCAAGCTCTGATCATATTTCTTTATTT 379
QY 362 TGATTTGGGANAANAATACTGTTTCTGATGACATGAATGCAAAATTTTGAATTTTAA 421
DB 380 TGATTTGGGANAANAATACTGTTTCTGATGACATGAATGCAAAATTTTGAATTTTAA 439
QY 422 TCTCACTAATTTTAANAACATATGAGAAATGATTAATGACATGAGTGACAACACTAA 481
DB 440 TCTCACTAATTTTAAGAACATATGAGAAATGATTAATGACATGAGTGACAACACTAA 499
QY 482 TTTACTGGCCAGCTGTGGCATGTGTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAA 541
DB 500 TTTACTGGCCAGCTGTGGCATGTGTTCTTACTTAGTTCTCCCAA-GGAAAACTCTTAA 558
QY 542 ATTGAATCTTCAGCAGATAATATCTTAATAATACTTTGTAAGCAAAACAAGCTTTT 601

|||||
DB 559 ACTGAATCTTCAGCAGATAATATCTTAATAATACTTTGTAAGCAAAACAAGCTTTT 618
QY 602 GTTTACATAGTTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACC 661
DB 619 GTTTACATAGTTCTTT-GGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACC 677
QY 662 CCAGACCATAATTACCATATTAACTTTGTTNCGACAGAGTTGTTGCCAAT 711
DB 678 CCAGACCATAATTACCATATTAACTTTGTTNATGACACAGTTGATGCAATT 727
RESULT 2
BM997078 761 bp mRNA linear EST 26-MAR-2002
LOCUS
DEFINITION
UI-H-ED0-axo-f-03-0-UI.s1 NCI_CGAP_ED0 Homo sapiens cDNA clone
IMAGE:5831426 3', mRNA sequence.
ACCESSION
BM997078 GI:19721979
VERSION
BM997078.1 GI:19721979
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 761)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 316-342, >AT rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
location/Qualifiers
1. 761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5831426"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED0"
/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI_CGAP_ED0 is a cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_LIB=UI-H-ED0
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"
BASE COUNT 269 a 108 c 98 g 285 t 1 others
ORIGIN
Query Match 94.2%; Score 672.8; DB 12; Length 761;
Best Local Similarity 98.2%; Pred. No. 7.4e-70;

[illegible]

Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
94.1%; Score 671.8; DB 9; Length 821;	96.9%; Pred. No. 9.4e-70;	690;	0;	21;	1;	1;
2	GCCAGAAAAAGTATTTTAATTTCTATTTAAACATCTTCTCAAGCATTTATTCCT	1	GCCAGAAAAAGTATTTTAATTTCTATTTAAACATCTTCTCAAGCATTTATTCCT	60		
62	ATATCTCACTGAATTTTAAGAATAACATTAGTATTAGAAAAAAGTAA	61	ATATCTCACTGAATTTTAAGAATAACATTAGTATTAGAAAAAAGTAA	120		
61	ATATCTCACTGAATTTTAAGAATAACATTAGTATTAGAAAAAAGTAA	120	ATATCTCACTGAATTTTAAGAATAACATTAGTATTAGAAAAAAGTAA	120		
122	TGCAGATAATTAACCTACATGAAAAAGAAAAATTAAACAAAGGACTGAAAGCTTATA	181	TGCAGATAATTAACCTACATGAAAAAGAAAAATTAAACAAAGGACTGAAAGCTTATA	181		
121	TGCAGATAATTAACCTACATGAAAAAGAAAAATTAAACAAAGGACTGAAAGCTTATA	180	TGCAGATAATTAACCTACATGAAAAAGAAAAATTAAACAAAGGACTGAAAGCTTATA	180		
182	AATTGAATGAGATTATAATTGAAAACTGCATCGAAAGCAAACTTATTGTTCAATTA	241	AATTGAATGAGATTATAATTGAAAACTGCATCGAAAGCAAACTTATTGTTCAATTA	240		
181	AATTGAATGAGATTATAATTGAAAACTGCATCGAAAGCAAACTTATTGTTCAATTA	240	AATTGAATGAGATTATAATTGAAAACTGCATCGAAAGCAAACTTATTGTTCAATTA	240		
242	TNCTTAATGATGGTGGTTTATGACTAATACACTGATTTTCAAGAAAGAAACCACTGTTA	301	TNCTTAATGATGGTGGTTTATGACTAATACACTGATTTTCAAGAAAGAAACCACTGTTA	301		
241	TTCTTAATGATGGTGGTTTATGACTAATACACTGATTTTCAATAGAAACCACTGTTA	300	TTCTTAATGATGGTGGTTTATGACTAATACACTGATTTTCAATAGAAACCACTGTTA	300		
302	AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAATT	361	AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAATT	361		
301	AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAATT	360	AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAATT	360		
362	TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA	421	TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA	421		
361	TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA	420	TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA	420		
422	TCTCACTAATTTTAAANAATACTGTTGAGAAATGATTAATGACATGAAGTGACAACTAA	481	TCTCACTAATTTTAAANAATACTGTTGAGAAATGATTAATGACATGAAGTGACAACTAA	481		
421	TCTCACTAATTTTAAANAATACTGTTGAGAAATGATTAATGACATGAAGTGACAACTAA	480	TCTCACTAATTTTAAANAATACTGTTGAGAAATGATTAATGACATGAAGTGACAACTAA	480		
482	TTACTGGCCAGCTGTGGCATGTGTTCTTACTAGTTCTCCCAAGGAAACTCTTAA	541	TTACTGGCCAGCTGTGGCATGTGTTCTTACTAGTTCTCCCAAGGAAACTCTTAA	541		
481	TTACTGGCCAGCTGTGGCATGTGTTCTTACTAGTTCTCCCAAGGAAACTCTTAA	539	TTACTGGCCAGCTGTGGCATGTGTTCTTACTAGTTCTCCCAAGGAAACTCTTAA	539		
542	ATTGAATCTTCAGCAGAAATACTCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT	601	ATTGAATCTTCAGCAGAAATACTCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT	601		
540	ACTGAATCTTCAGCAGAAATACTCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT	599	ACTGAATCTTCAGCAGAAATACTCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT	599		

QY 602 GTTACATAGTCTTGGGATTTTACTGTCCATAATTTATCTGAACCTCAATTTTACC 661
| | | | |
Db 600 GTTACATAGNTCTTTGGATTTTACTGTCTTAATTTTATTCTGAAACTCAATTTACC 659
| | | | |
QY 662 CCAGACCATAATTACCATTAATTAAGTTGTTGACAGTTGTTGCCAATTC 713
| | | | |
Db 660 CCAGACCATAATTACCATTAACNTTGTATGACAGTTGTATGCAATTC 711
| | | | |
RESULT 4 770 bp mRNA linear EST 21-DEC-1999
AI826287 WK33f07.x1 NCI CGAP Pr22 Homo sapiens cDNA clone IMAGE:2417221 3'
LOCUS similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR
DEFINITION (HUMAN); contains element MER28 repetitive element ; mRNA sequence.
AI826287
ACCESSION AI826287.1 GI:5446958
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 770)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1454 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 404.
Location/Qualifiers
1. 770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2417221"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Pr22"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; 1st strand cDNA was prepared
from normal prostate bulk tissue, and was then primed with
a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 271 a 111 c 99 g 285 t 4 others
ORIGIN
Query Match 93.9%; Score 670.4; DB 9; Length 770;
Best Local Similarity 97.8%; Pred. No. 1.4e-69;
Matches 696; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 122 TGCAGATATTAACCTTACATGAAAAAGGAAATTTATACAAAGGACTGACAAGCTTATA 181
| | | | |
Db 145 TGCAGATATTAACCTTACATGAAAAAGGAAATTTATACAAAGGACTGACAAGCTTATA 204
| | | | |
QY 182 AATTGAATGAGATTATAATTGAAAAGCTGATCTGAAAAGCAACTTTATTGTTCAATTA 241
| | | | |
Db 205 AATTGAATGAGATTATAATTGAAAAGCTGATCTGAAAAGCAACTTTATTGTTCAATTA 264
| | | | |
QY 242 TNCCTAATGATGCTGTTTATGACTAATACTGATTTTTCAGAAGGAAACCATGTTA 301
| | | | |
Db 265 TTCTTAATGATGCTGTTTATGACTAATACTGATTTTTCAGAAGGAAACCATGTTA 324
| | | | |
QY 302 AAAATATTTTATTTTAAATAAGCCGTGTTCAAGCTCTGATCATATTTCTTTATTT 361
| | | | |
Db 325 AAAATATTTTATTTTAAATAAGCCGTGTTCAAGCTCTGATCATATTTCTTTATTT 384
| | | | |
QY 362 TGATTTGGGAAANAATACTCTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 421
| | | | |
Db 385 TGATTTGGGAGAAATACTCTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 444
| | | | |
QY 422 TCTCACTAATTTTAAANAATACTATGAGAAATTGATTAATGACATGAGTGACACACTAA 481
| | | | |
Db 445 TCTCACTAATTTTAAAGAACTATGAGAAATTGATTAATGACATGAGTGACACACTAA 504
| | | | |
QY 482 TTAATGCGCAGCTGTTGCAATGTTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAA 541
| | | | |
Db 505 TTAATGCGCAGCTGTTGCAATGTTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAA 563
| | | | |
QY 542 ATTGAATCTTCAGAGATAATATCTTAATAATACTTTGTAAGCAAAACAAAGCTTTT 601
| | | | |
Db 564 ACTGAATCTTCAGAGATAATATCTTAATAATACTTTGTAAGCAAAACAAAGCTTTT 623
| | | | |
QY 602 GTTACATAGTCTTGGGATTTTACTGTTCTTAATTTATCTGAAAACCTCAATTTTACC 661
| | | | |
Db 624 GTTACATAGTCTTTT-GGATTNACTGTTCTTAATTTATCTGAAAACCTCAATTTTACC 682
| | | | |
QY 662 CCAGACCATAATTACCATATTAATTAAGTTGTTGACAGAGTTGTTGCCAATTC 713
| | | | |
Db 683 CCAGACCATAATTACCATATTAATTAAGTTGTTGACAGAGTTGTTGCCAATTC 734
| | | | |
RESULT 5 756 bp mRNA linear EST 20-FEB-2003
BM985376
LOCUS UI-CF-EC1-acg-p-23-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
DEFINITION UI-CF-EC1-acg-p-23-0-UI 3', mRNA sequence.
BM985376
ACCESSION BM985376.1 GI:19611803
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 756)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research

Genetics (www.reegen.com) or from Open Biosystems (www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 316-342, >AT_rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES

source

Location/Qualifiers
1..756
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acg-p-23-0-UI"
/issue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker, Site_1: EcoR I; Site_2: Not I; UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AAGTGCTTAC.
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells tissue nos 369-371 and 380-383
TAG_SEQ=AAGTGCTTAC"

BASE COUNT 268 a 107 c 97 g 283 t 1 others
ORIGIN

Query Match 93.8%; Score 669.6; DB 12; Length 756;

Best Local Similarity 97.9%; Pred. No. 1.8e-69;

Matches 697; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

OY 2 GCCAGAAAAAGTTATTTTATTTCTATTAACATCTTCTCAAAGCATTTATCTCT 61
DB 19 GCCAGAAAAAGTTATTTTATTTCTATTAACATCTTCTCAAAGCATTTATCTCT 78
OY 62 ATATCTCACTGAATTTAAGAATAACATAGTATTAGAAAACTAGAAAAAGATAAA 121
DB 79 ATATCTCACTGAATTTAAGAATAACATAGTATTAGAAAACTAGAAAAAGATAAA 138
OY 122 TGCAGATAATTAACCTACATGAAAAAGAAAAATTATAACAAGGAGCTGAGAACGTTATA 181
DB 139 TGCAGATAATTAACCTACATGAAAAAGAAAAATTATAACAAGGAGCTGAGAACGTTATA 198
OY 182 AATTGAATGAGATTATAATTGAAAACCTGCATCTGAAAGCAACTTATTGTTCAATTA 241
DB 199 AATTGAATGAGATTATAATTGAAAACCTGCATCTGAAAGCAACTTATTGTTCAATTA 258
OY 242 TNCTTAATGATGGTGTATTATGACTAATACACTGATTTTCAAGAAGGAAACCCATGTTA 301
DB 259 TTCTTAATGATGGTGTATTATGACTAATACACTGATTTTCAATAAGGAAACCCATGTTA 318
OY 302 AAATATATTTTATTTTAAAAAATAGCCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 361
DB 319 AAATATATTTTATTTTAAAAAATAGCCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 378
OY 362 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 421
DB 379 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 438
OY 422 TCTCACTAATTTTANAACATATGAGAATGATTAATGACATGAGTGACACAACACTAA 481
DB 439 TCTCACTAATTTTANAACATATGAGAATGATTAATGACATGAGTGACACAACACTAA 498

OY 482 TTACTGGCCAGCTGTGGCATTTGTGTTCTTACTAGTTCCTCCCAAGGAAAACTCTAA 541
DB 499 TTACTGGCCAGCTGTGGCATTTGTGTTCTTACTAGTTCCTCCCAAGGAAAACTCTAA 557
OY 542 ATTGAATCTTCAGCAGATAATCCCTTAATATACCTTTGTAGCAAAACAAAGCTTTT 601
DB 558 ACTGAATCTTCAGCAGATAATCCCTTAATATACCTTTGTAGCAAAACAAAGCTTTT 617
OY 602 GTTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATTCGAACTCAATTTTACC 661
DB 618 GTTTACATAGTCTTTT-GGATTTTACTGTTCTTAATTTTATTCGAACTCAATTTTACC 676
OY 662 CCAGACCAATATTACCATATTACTTTGTNTGCACAGTTGTTGCCAATTC 713
DB 677 CCAGACCAATATTACCATATTACTTTGTATGACACAGTTGTATGCAATTCC 728

RESULT 6

BQ045161

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.reegen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 316-342, >AT_rich#Low_complexity (matched complement)

Seq primer: M13 FORWARD

POLYA=yes.

FEATURES

source

Location/Qualifiers
1..738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-aei-a-17-0-UI"
/issue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker, Site_1: EcoR I; Site_2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGCT.

TAG_LIB=UI-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTGCTCAGCT"
BASE COUNT 260 a 102 c 97 g 278 t 1 others
ORIGIN

Query Match 93.6%; Score 668.4; DB 12; Length 738;
Best Local Similarity 98.3%; Pred. No. 2.5e-69;
Matches 694; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

Qy 2 GCCAGAAAAAGTTATTTTAACTTTCTATTAACATCTCTCTCAAGCATTTATTCCT 61
Db 19 GCCAGAAAAAGTTATTTTAACTTTCTATTAACATCTCTCTCAAGCATTTATTCCT 78
Qy 62 ATATCTCACTGAATTTTAAAGAAATACATTAAGTAAAGAAAGTAAAGATATAA 121
Db 79 ATATCTCACTGAATTTTAAAGAAATACATTAAGTAAAGAAAGTAAAGATATAA 138
Qy 122 TGCAGATAATTAACCTTACATGAAAAAGAAATTAACAAGGAGCTGAGACGTTATA 181
Db 139 TGCAGATAATTAACCTTACATGAAAAAGAAATTAACAAGGAGCTGAGACGTTATA 198
Qy 182 AATGAATGAGATTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTA 241
Db 199 AATGAATGAGATTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTA 258
Qy 242 TNCCTTAATGAGTGTGTTTATGACTAATACATGATTTTTCAGAGAAAGAAACCATGTTA 301
Db 259 TTCTTAATGAGTGTGTTTATGACTAATACATGATTTTTCAGAGAAAGAAACCATGTTA 318
Qy 302 AAAATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATTT 361
Db 319 AAAATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATTT 378
Qy 362 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATTTTAA 421
Db 379 TGATTTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATTTTAA 438
Qy 422 TCTCACTAATTTAANAACATATTGAGAAATGATTAATGACATGAGTGCAACACTAA 481
Db 439 TCTCACTAATTTAAGAACAATGAGAAATGATTAATGACATGAGTGCAACACTAA 498
Qy 482 TTTACTGGCCAGCTGTGGCATTGTGTTCTTACTAGTCTCCCAAGGAAACCTCTTAA 541
Db 499 TTTACTGGCCAGCTGTGGCATTGTGTTCTTACTAGTCTCCCAAGGAAACCTCTTAA 557
Qy 542 ATTGAATCTTCAGAGAAATAATCCTTAATAATACTTTGTAAGCAAAACAAAGCTTTT 601
Db 558 ACTGAATCTTCAGAGAAATAATCCTTAATAATACTTTGTAAGCAAAACAAAGCTTTT 617
Qy 602 GTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACC 661
Db 618 GTTACATAGTCTTT-GGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACC 676
Qy 662 CCAGACCATAATTAACATTAATTAATTTGTTGCAAGTGTGTC 707
Db 677 CCAGACCATAATTAACATTAATTAATTTGTTGCAAGTGTGTC 722

RESULT 7
AW052045 786 bp mRNA linear EST 20-SEP-1999
LOCUS ,
DEFINITION wx25e05.x1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:2544704 3'
similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR
(HUMAN);, mRNA sequence.

ACCESSION AW052045
VERSION AW052045.1 GI:5914404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 786)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL
COMMENT
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 458.
Location/Qualifiers

FEATURES
source
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2544704"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid1"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 290 a 116 c 109 g 270 t 1 others
ORIGIN

Query Match 93.4%; Score 667; DB 9; Length 786;
Best Local Similarity 97.6%; Pred. No. 3.5e-69;
Matches 695; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

Qy 2 GCCAGAAAAAGTTATTTTAACTTTCTATTAACATCTCTCAAGCATTTATTCCT 61
Db 2 GCCAGAAAAAGTTATTTTAACTTTCTATTAACATCTCTCGGAAAGCATTTATTCCT 61
Qy 62 ATATCTCACTGAATTTTAAAGAAATACATTAATTAAGAAACTGAGAAAGATATAA 121
Db 62 ATATCTCACTGAATTTTAAAGAAATACATTAATTAAGAAACTGAGAAAGATATAA 121
Qy 122 TGCAGATAATTAACCTTACATGAAAAAGAAATTTATAACAAGGAGCTGAGAACGTTATA 181
Db 122 TGCAGATAATTAACCTTACATGAAAAAGAAATTTATAACAAGGAGCTGAGAACGTTATA 181
Qy 182 AATGAATGAGATTATAATTGAAAACTGCATCTGAAAGCAAACTTATTTGTTCAATTA 241
Db 182 AATGAATGAGATTATAATTGAAAACTGCATCTGAAAGCAAACTTATTTGTTCAATTA 241
Qy 242 TNCCTTAATGAGTGTGTTTATGACTAATACTGATTTTTCAGAGAAAGAAACCATGTTA 301
Db 242 TNCCTTAATGAGTGTGTTTATGACTAATACTGATTTTTCAGAGAAAGAAACCATGTTA 301
Qy 302 AAAATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTT 361
Db 302 AAAATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTT 361

OY		362	TGATTTGGGAANNAATACTGTTTCTGTATAGCATGAAATGCAAAATTTTAGATTTTAA	421
Db		362	TGATTTGGGAAGAAATACTGTTTCTGTATAGCATGAAATGCAAAATTTTAGATTTTAA	421
OY		422	TCTCACTAATTTTAANAACATATGAGAATTGATTAATGACATGAAGTGACAACACTAA	481
Db		422	TCTCACTAATTTTAAGAACATATGAGAATTGATTAATGACATGAAGTGACAACACTAA	481
OY		482	TTACTGGCCAGCTGTGGCATTGTGTTTCTTACTTAGTTCGCCAAGGAAAACTTTAA	541
Db		482	TTACTGGCCAGCTGTGGCATTGTGTTTCTTACTTAGTTCGCCAA-GGAAAACCTTTAA	540
OY		542	ATTGAATCTTCAGCAGATAATCCTTAATATACCTTTGTAAGCAAACAAAGCTTTTTT	601
Db		541	ACTGAATCTTCAGCAGATAATCCTTAATATACCTTTGTAAGCAAACAAAGCTTTTTT	600
OY		602	GTTTACATAGTCTTTGGGATTTTACTGTTCTCCTAATTTTAATCTGAAACTCAATTTTACC	661
Db		601	GTTTACATAGTCTTTT-GGATTTTACTGTTCTCCTAATTAATCTGAAACTCAATTTTACC	659
OY		662	CCAGACCATATATCACCATATTACTTTGTNTGACACAGTTGTTGCCAATTC	713
Db		660	CCAGACCATATATCACCATATTACTTTGTATATGACACAGTTGTAATGC AATTNC	711
RESULT 8				
LOCUS	BQ008197	751 bp	mRNA	linear EST 26-MAR-2002
DEFINITION	UI-H-ED1-ayc-h-19-0-UI.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone IMAGE:5836866 3', mRNA sequence.			
ACCESSION	BQ008197			
VERSION	BQ008197.1	GI:19733097		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 751) NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished			
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs-rc@mail.nih.gov Tissue Procurement: Dr. Jose Mercuende cDNA Library preparation by: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.lnll.gov The following repetitive elements were found in this cDNA sequence: 319-345, >AT-rich#low_complexity (matched complement) Seq primer: M13 FORWARD PolyA=Yes.			
FEATURES	Location/Qualifiers			
Source	1..751			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:5836866"			
	/tissue_type="Chondrosarcoma"			
	/dev_stage="Adult"			
	/lab_host="DH10B (Life Technologies)"			
	/clone_1lb="NCI_CGAP_ED1"			
	/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP_ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line CSS. The library was constructed according to Bonaldde , Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and			

BASE COUNT	266 a	104 c	98 g	282 t	1 others
Query Match	93.1%;	Score 664.4;	DB 12;	Length 751;	
Best Local Similarity	97.0%;	Pred. No. 7.2e-69;			
Matches 685;	Conservative 0;	Mismatches 20;	Indels 1;	Gaps 1;	
<p>cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)₁₈ tail. The TAG_LIB=UI-H-ED1 TAG_TISSUE=chondrosarcoma TAG_SEQ=CGTCAAGGCT"</p>					
Db	2	GCCAGAAAAAGTATTTTAAATTTCTATTAACAATTCTTCTCAAGCATTTATTTA	61		
Db	22	GCCAGAAAAAGTATTTTAAATTTCTATTAACAATTCTTCTCAAGCATTTATTTA	81		
Db	62	ATATCTCACTGAATTTTAAAGAAATACATAGTATTAGAAAACTAGGAAAAAGAT	121		
Db	82	ATATCTCACTGAATTTTAAAGAAATACATAGTATTAGAAAACTAGGAAAAAGAT	141		
Db	122	TGCAGATAATTAACCTACATGAAAAAGAAAAATTAACAAGGACTGAGAACGTTA	181		
Db	142	TGCAGATAATTAACCTACATGAAAAAGAAAAATTAACAAGGACTGAGAACGTTA	201		
Db	182	AATGAATGAGATTAATTTTGAAGAACTGCATCTGAAGCAAACTTATTTGTA	241		
Db	202	AATGAATGAGATTAATTTTGAAGAACTGCATCTGAAGCAAACTTATTTGTA	261		
Db	242	TNCTTAATGATGGTGTATTTATGACTAATACACTGATTTTCAAGAAAGAA	301		
Db	262	TCTTAATGATGGTGTATTTATGACTAATACACTGATTTTCAATAGAAACC	321		
Db	302	AAAATATTTTATTTTAAATAAAGCCGTGTGTTCAAGCTCTGATCATATTT	361		
Db	322	AAAATATTTTATTTTAAATAAAGCCGTGTGTTCAAGCTCTGATCATATTT	381		
Db	362	TGATTTGGGAAANAATACTCTGTTCTGATAGCATGAATGCCAAATTTT	421		
Db	382	TGATTTGGGAAANAATACTCTGTTCTGATAGCATGAATGCCAAATTTT	441		
Db	422	TCTCACTAATTTTAANAATACTATTGAGAAATTGATTAATGACATGAAGT	481		
Db	442	TCTCACTAATTTTAANAATACTATTGAGAAATTGATTAATGACATGAAGT	501		
Db	482	TTACTGGCCAGCTGTGGCAATGTGTTCTTACTTACTGTTCTCCCAAGGAA	541		
Db	502	TTACTGGCCAGCTGTGGCAATGTGTTCTTACTTACTGTTCTCCCAAGGAA	560		
Db	542	ATTGAATCTTCAGCAGATAATCCTTAATATACCTTTGTAAGCAAAACAA	601		
Db	561	ACTGAATCTTCAGCAGATAATCCTTAATATACCTTTGTAAGCAAAACAA	620		
Db	602	GTTTACATAGTCTTTGGGATTTTACTGTCTCTAATTTTATTTCTGAAACT	661		
Db	621	TGTTTACATAGTCTTTGGGATTTTACTGTCTCTAATTTTATTTCTGAAACT	680		
Db	662	CCAGACCATTAATTAACCAATTAACCTTTGTTNGCAGAGTGTGTTGC	707		
Db	681	CCAGACCATTAATTAACCAATTAACCTTTGTTNGCAGAGTGTGTTGC	726		
RESULT 9	BU619112	759 bp	mRNA	linear	EST 23-SEP-2002
LOCUS	UI-H-FH1-bfm-1-10-0-UI.s1	NCI	CGAP	FH1 Homo sapiens	cDNA clone
DEFINITION	UI-H-FH1-bfm-1-10-0-UI 3'				mRNA sequence.
ACCESSION	BU619112				
VERSION	BU619112.1	GI:23285327			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 759)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 316-342, >AT rich#low_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES

Source
 Location/Qualifiers
 1..759
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FH1-bfm-1-10-0-UI"
 /tissue_type="Cell line"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_FH1"
 /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; NCI CGAP_FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGGC. The cell line was provided by Dr. James Martin from the University of Iowa."
 TAG_LIB=UI-H-FH1
 TAG_TISSUE=Human Chondrosarcoma Cell Line C58 - Grade 1
 Chondrosarcoma
 TAG_SEQ=AGAATCCGGC"
 BASE COUNT 269 a 97 g 286 t
 ORIGIN

Query Match 93.1%; Score 664.4; DB 13; Length 759;
 Best Local Similarity 97.0%; Pred. No. 7.1e-69;
 Matches 685; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 2 GCCAGAAAAAGTTAATTTCTATTAACATCTTCTCAAGCATATTATTCCT 61
 Db 19 GCCAGAAAAAGTTAATTTCTATTAACATCTTCTCAAGCATATTATTCCT 78
 QY 62 ATATCTCACTGAATTTTAAGAAATACATTAAGAAAACTAGAAAAAGATAAA 121
 Db 79 ATATCTCACTGAATTTTAAGAAATACATTAAGAAAACTAGAAAAAGATAAA 138
 QY 122 TGCAGATAATTAACCTACATGAAAAAGAAATTAACAAAGAGCTGAGAAAGTTATA 181
 Db 139 TGCAGATAATTAACCTACATGAAAAAGAAATTAACAAAGAGCTGAGAAAGTTATA 198
 QY 182 AATTGAATGAGATTATTAATTTGAAAACTGCAAGCAAACTTTATGTTCAATTA 241
 Db 199 AATTGAATGAGATTATTAATTTGAAAACTGCAAGCAAACTTTATGTTCAATTA 258

QY 242 TNCCTAATGATGGTGTATTGACTAATACACTGATTTTCAAGAGAAACCCAGTTA 301
 Db 259 TCTTAATGATGGTGTATTGACTAATACACTGATTTTCAATAGAAACCCAGTTA 318
 QY 302 AAATATTTTATTTTAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATT 361
 Db 319 AAATATTTTATTTTAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATT 378
 QY 362 TGAATTTGGGAANAATACCTGTTTCTGATAGCATGAAATGCAAAATTTTAGATT 421
 Db 379 TGAATTTGGGAANAATACCTGTTTCTGATAGCATGAAATGCAAAATTTTAGATT 438
 QY 422 TCTCACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAGTGACACACTAA 481
 Db 439 TCTCACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAGTGACACACTAA 498
 QY 482 TTAAGTGGCAGCTGTGGCATTTGTTTCTTACTAGTTCTCCCAAGGAAACTCTTAA 541
 Db 499 TTAAGTGGCAGCTGTGGCATTTGTTTCTTACTAGTTCTCCCAAGGAAACTCTTAA 557
 QY 542 ATTGAATCTTCAGACAGATAATCCTTAATATACCTTTGTAAGCAAAACAAGCTTTT 601
 Db 558 ACTGAATCTTCAGACAGATAATCCTTAATATACCTTTGTAAGCAAAACAAGCTTTT 617
 QY 602 GTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACC 661
 Db 618 GTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACC 677
 QY 662 CCAGACCATATTAACATATTAACTTTGTTGTCAGACAGTTGTTGC 707
 Db 678 CCAGACCATATTAACATATTAACTTTGTTGTCAGACAGTTGTTGC 723

RESULT 10
 LOCUS BM968746 721 bp mRNA linear EST 20-FEB-2003
 DEFINITION UI-CF-DU1-aam-a-03-0-UI.s1 UI-CF-DU1 Homo sapiens cDNA clone
 UI-CF-DU1-aam-a-03-0-UI 3', mRNA sequence.
 ACCESSION BM968746
 VERSION BM968746
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 721)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL MEDLINE
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 The following repetitive elements were found in this cDNA
 sequence: 315-341, >AT rich#low_complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES
 source
 Location/Qualifiers
 1..721
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-aam-a-03-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

BASE COUNT      257 a      97 c      91 g      276 t
ORIGIN

Query Match      93.0%; Score 664.2; DB 12; Length 721;
Best Local Similarity 98.0%; Pred. No. 7.7e-69;
Matches 691; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY      3 CCAGAAAAAGTATTTTCTATTAACATCTCTCAAGCATTTTATCCTA 62
      19 CCAGAAAAAGTATTTTCTATTAACATCTCTCAAGCATTTTATCCTA 78
QY      63 TATCTCACTGAATTTAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATAAT 122
      79 TATCTCACTGAATTTAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATAAT 138
Db
QY      123 GCAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAGACTGAGAACGTTATA 182
      139 GCAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAGACTGAGAACGTTATA 198
QY      183 ATTGAATGAGATTATAATTGAAAACTGCATCTGAAGCAAACTTATTGTTCAATTAT 242
      199 ATTGAATGAGATTATAATTGAAAACTGCATCTGAAGCAAACTTATTGTTCAATTAT 258
Db
QY      243 NCTTAATGATGTTTATGACTAATACACTGATTTTCAAGGAAGAAACCCATGTTAA 302
      259 TCTTAATGATGTTTATGACTAATACACTGATTTTCAAGGAAGAAACCCATGTTAA 318
Db
QY      303 AAATATTTTATTTAAATAAGCCCTGTGTCAAGCTCTGATCATATTTCTTTATTTT 362
      319 AAATATTTTATTTAAATAAGCCCTGTGTCAAGCTCTGATCATATTTCTTTATTTT 378
QY      363 GATTGGGAANAATACTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTAAAT 422
      379 GATTGGGAANAATACTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTAAAT 438
Db
QY      423 CTCACATAATTTAANAACATAAGAAATTGATTAATGACATGAAGTGACAACACTAAT 482
      439 CTCACATAATTTAAGACATAAGAAATTGATTAATGACATGAAGTGACAACACTAAT 498
Db
QY      483 TACTGGCCAGCTGTGGCATTTGTTTCTTACTTAGTCTCCCAAGGAAACTCTTAA 542
      499 TACTGGCCAGCTGTGGCATTTGTTTCTTACTTAGTCTCCCAAGGAAACTCTTAA 557
QY      543 TTGAATCTTCAGCAGAAATATCTTAAATATACCTTTGTAAGCAAAACAAAGCTTTT 602
      558 CTGAATCTTCAGCAGAAATATCTTAAATATACCTTTGTAAGCAAAACAAAGCTTTT 617
QY      603 TTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACC 662
      |||||||
```

```

Db      618 TTTACATAGTCTTT-GGATTTTACTGTTCCTAATTTATCTGAAACTCAATTTTACC 676
QY      663 CAGACCATAATTACATATTAATTTGTTTNGCAGAGTTGTC 707
      |||||
Db      677 CAGACCAATATTTACATATTAATTTGTTGTAAGCAGATTGTATGC 721

RESULT 11
CA424156
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-48, >AT-rich#low_complexity (matched complement)
314-340, >AT-rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source
Location/Qualifiers
1..734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-bdv-m-12-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FEI"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site_1: EcoR I; Site_2: Not
I; NCI_CGAP_FEI is a normalized cDNA library derived from
a pool of mRNA obtained from 3 cell lines from grade II
chondrosarcoma tissues. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGCTACGAGC. The cell
lines were provided by Dr James Martin from the University
of Iowa.
TAG_LIB=UI-H-FEI
TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG_SEQ=CGCTACGAGC"

BASE COUNT      261 a      101 c      94 g      277 t      1 others
ORIGIN

Query Match      92.9%; Score 663.2; DB 14; Length 734;
Best Local Similarity 98.0%; Pred. No. 1e-68;
Matches 690; Conservative 0; Mismatches 12; Indels 2; Gaps 2;
```

QY 4 CAGAAAAAGTATTATTATTCTATTAAACATTCTTCTCAAGCATTATTATCTAT 63
| | | | |
Db 19 CAGAAAAAGTATTATTATTCTATTAAACATTCTTCTCAAGCATTATTATCTAT 78
| | | | |
QY 64 ATCTCACTGAATTTTAAAGAAATACATAGTATTAGAAAACTAGAAAAAGATAATG 123
| | | | |
Db 79 ATCTCACTGAATTTTAAAGAAATACATAGTATTAGAAAACTAGAAAAAGATAATG 138
| | | | |
QY 124 CAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAGGACTGGAACGTTATAA 183
| | | | |
Db 139 CAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAGGACTGGAACGTTATAA 198
| | | | |
QY 184 TTGAATGAGATTATATTGAAAACTGCATCTGAAAGCAACTTTATTGTTCAATTATN 243
| | | | |
Db 199 TTGAATGAGATTATATTGAAAACTGCATCTGAAAGCAACTTTATTGTTCAATTAT 258
| | | | |
QY 244 CTTAATGATGTTGTTTATGACTAATACACTGATTTTTCAAGAGAAACCCATGTTAA 303
| | | | |
Db 259 CTTAATGATGTTGTTTATGACTAATACACTGATTTTTCAATGAAGAAACCCATGTTAA 318
| | | | |
QY 304 AATATTTTATTAAAAAATAAGCCTGTGTTCAAGCTGATCATATTCTTTATTTTG 363
| | | | |
Db 319 AATATTTTATTAAAAAATAAGCCTGTGTTCAAGCTGATCATATTCTTTATTTTG 378
| | | | |
QY 364 ATTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATC 423
| | | | |
Db 379 ATTGGGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATC 438
| | | | |
QY 424 TCACATAATTTTAAANAACATTTGAGAAATGATTAATGACATGAAGTGACACACTAAT 483
| | | | |
Db 439 TCACATAATTTTAAAGACATTTGAGAAATGATTAATGACATGAAGTGACACACTAAT 498
| | | | |
QY 484 ACTGGCCAGCTGTGGCATGTGTTCTTACTTACTTCTCCCAAGGAAAACTTTAAAT 543
| | | | |
Db 499 ACTGGCCAGCTGTGGCATGTGTTCTTACTTACTTCTCCCAAGGAAAACTTTAAAC 557
| | | | |
QY 544 TGAATCTTCAGCAGAAATATCTTAAATATATCTTTGTAAGCAAAACAAAGCTTTTGT 603
| | | | |
Db 558 TGAATCTTCAGCAGAAATATCTTAAATATATCTTTGTAAGCAAAACAAAGCTTTTGT 617
| | | | |
QY 604 TTACATAGTCTTTGGGATTTTACTGTTCTTAAATTTTCTGAAACTCAATTTTACCC 663
| | | | |
Db 618 TTACATAGTCTTTT-GGATTTTACTGTTCTTAAATTTTCTGAAACTCAATTTTACCC 676
| | | | |
QY 664 AGACCATATATACCATATTAATCTTTGTTNGACAGTTGTTGC 707
| | | | |
Db 677 CAGACATATATACCATATTAATCTTTGTTNGACAGTTGTTATGC 720
| | | | |

RESULT 12
BU627147 762 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-FG0-bct-i-12-0-UI.s1 NCI_CGAP_EN1_2 Homo sapiens cDNA clone
DEFINITION UI-H-FG0-bct-i-12-0-UI 3', mRNA sequence.
ACCESSION BU627147
VERSION BU627147.1 GI:232933361
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 762)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-43, >AT_rich#Low_complexity (matched complement)
309-335, >AT_rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers
1..762
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FG0-bct-i-12-0-UI"
/tissue_type="Enchondroma cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_EN1_2"
/note="Organ: Bone; Vector: pTR73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
directionally into pTR73-Pac vector. The oligonucleotide
tissue(s): Enchondroma cell line (2 cell lines). The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pTR73-Pac vector. The oligonucleotide
site and the (dT)18 tail. The sequence tag for this
library is CCGGTCACTC. The cell lines was provided by Dr
James Martin from University of Iowa.
TAG_LIB=UI-H-FG0
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG_SEQ=CGGTCACTC"
BASE COUNT 269 a 109 c 97 g 286 t 1 others
ORIGIN

Query Match 92.6%; Score 661.4; DB 13; Length 762;
Best Local Similarity 98.3%; Pred. No. 1.6e-68;
Matches 687; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 9 AAAGTTATTTAATTTTCTATTAAACATCTCTCAAGCATTAATTTATCTATCTC 68
| | | | |
Db 19 AAAGTTATTTAATTTTCTATTAAACATCTCTCTCAAGCATTAATTTATCTATCTC 78
| | | | |
QY 69 ACTGAATTTTGAAGAAATACATTAAGTAAAGCAAAAGATAAATGCAGAT 128
| | | | |
Db 79 ACTGAATTTTGAAGAAATACATTAAGTAAAGCAAAAGATAAATGCAGAT 138
| | | | |
QY 129 AATTAACTTACATGAAAAAGAAATTTATACAAAGGACTGAGACGTTATAATTGAA 188
| | | | |
Db 139 AATTAACTTACATGAAAAAGAAATTTATACAAAGGACTGAGACGTTATAATTGAA 198
| | | | |
QY 189 ATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATNCTTA 248
| | | | |
Db 199 ATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATNCTTA 258
| | | | |
QY 249 TGATGGTGTTTTATGACTAATACACTGATTTTTCAGAGAGAAACCCATGTTAAATAAT 308
| | | | |
Db 259 TGATGGTGTTTTATGACTAATACACTGATTTTTCAGAGAGAAACCCATGTTAAATAAT 318
| | | | |
QY 309 TTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGATTG 368
| | | | |
Db 319 TTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGATTG 378
| | | | |
QY 369 GGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTATCTCACT 428
| | | | |
Db 379 GGAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTATCTCACT 438
| | | | |
QY 429 AATTTTAAANACTATTGAGAAATGATTAATGACATGAAGTGACAACAATAATTAATCTG 488
| | | | |
Db 439 AATTTTAAAGACTATTGAGAAATGATTAATGACATGAAGTGACAACAATAATTAATCTG 498
| | | | |

QY 489 CCAGCTGTGGCATGTGTTCTTACTTAGTTCCTCCGAAGGAAACTCTTAAATGAAT 548
|||||
Db 499 CCAGCTGTGGCATGTGTTCTTACTTAGTTCCTCCAA-GGAAAACTCTTAACTGAAT 557
QY 549 CTTGACGAGATAATCCTTAATATACTTTGTAGCAAAACAAAGCTTTTGTGTACA 608
|||||
Db 558 CTTGACGAGATAATCCTTAATATACTTTGTAGCAAAACAAAGCTTTTGTGTACA 617
QY 609 TAGTCTTTGGGATTTTACTGTCTCTAATTTTCTGAACCTCAATTTTACCCAGACC 668
|||||
Db 618 TAGTCTTT-GGATTTTACTGTCTCTAATTTTCTGAACCTCAATTTTACCCAGACC 676
QY 669 ATAATTACCATATTAACCTTTGTNTGCACAGTTGTTGC 707
|
Db 677 AATATTACCATATTAACCTTTGTATGACAGTTGTATGC 715

RESULT 13
BU933572/c 734 bp mRNA linear EST 18-OCT-2002

LOCUS BU933572 734 bp mRNA linear EST 18-OCT-2002
DEFINITION AGENCOURT 10507105 NIH_MGC_127 Homo sapiens CDNA clone
IMAGE:6698694 5', mRNA sequence.

ACCESSION BU933572
VERSION BU933572.1 GI:24122391
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 734)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LCM2985 row: f column: 06
High quality sequence stop: 526.

FEATURES
source
1. 734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6698694"
/issue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_id="NIH_MGC_127"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcggc);
Site 2: SfiI (ggcgcctggc); Double-stranded CDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGTATCAACGACGAGAGGCGCATTCGCGCGG-3' and
5'-ATTCTAGAGCGGAGGCGGCGGCGCATG-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 1-2 kb
size fraction (other fractions present in NIH_MGC_126 and
NIH_MGC_128). Library created in the laboratory of T.
Uedln, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
Library."

BASE COUNT 271 a 98 c 101 g 256 t 8 others
ORIGIN

Query Match 92.5%; Score 660.6; DB 13; Length 734;
Best Local Similarity 98.4%; Pred. No. 2e-68;
Matches 685; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 2 GCCAGAAAAAGTATTTTAAATTTTCTATTAACATTTCTCTCAAGCATTTATTCCT 61
|||||
Db 695 GCCAGAAAAAGTATTTTAAATTTTCTATTAACATTTCTCTCAAGCATTTATTCCT 636
QY 62 ATATCTCAGTGAATTTTAAAGAAATACATAGTATTAGAAAACTAGAAAAAGATAAA 121
|||||
Db 635 ATATCTCAGTGAATTTTAAAGAAATACATAGTATTAGAAAACTAGAAAAAGATAAA 576
QY 122 TGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGACGTATA 181
|||||
Db 575 TGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGACGTATA 516
QY 182 AATTGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTA 241
|||||
Db 515 AATTGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTA 456
QY 242 TNCCTAATGATGTTGTTTATGACTAATGACTGATTTTTCAGAGAGAAACCATGTTA 301
|||||
Db 455 TTCTTAATGATGTTGTTTATGACTAATGACTGATTTTTCAGAGAGAAACCATGTTA 396
QY 302 AAAATATTTTATTTTAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTT 361
|||||
Db 395 AAAATATTTTATTTTAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTT 336
QY 362 TGATTTGGGAAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA 421
|||||
Db 335 TGATTTGGGAAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA 276
QY 422 TCTCACTAATTTTAAANAATACTATGAGAAATGATTAATGACATGAAGTGCAACACTAA 481
|||||
Db 275 TCTCACTAATTTTAAANAATACTATGAGAAATGATTAATGACATGAAGTGCAACACTAA 216
QY 482 TTAGTGCCAGCTGTGGCATTTGTTCTTACTTACTTCTCCCAAGGAAAACTCTTAA 541
|||||
Db 215 TTAGTGCCAGCTGTGGCATTTGTTCTTACTTACTTCTCCCAAGGAAAACTCTTAA 157
QY 542 ATTGAATCTTCAGAGAGATAATCTTAATATCTTTGTAAGCAAAACAAAGCTTTT 601
|||||
Db 156 ACTGAATCTTCAGAGAGATAATCTTAATATCTTTGTAAGCAAAACAAAGCTTTT 97
QY 602 GTTTACATAGTTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACC 661
|||||
Db 96 GTTTACATAGTTCTTT-GGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACC 38
QY 662 CCAGACCATTAATTAACCATATTAACCTTTGTTNGCAC 697
|||||
Db 37 CCAGACCATTAATTAACCATATTAACCTTTGTAATGCAC 2

RESULT 14
BU619082 712 bp mRNA linear EST 23-SEP-2002
LOCUS BU619082
DEFINITION UI-H-FH1-bfm-f-06-0-UI.s1 NCI_CGAP_FH1 Homo sapiens CDNA clone
UI-H-FH1-bfm-f-06-0-UI 3', mRNA sequence.
ACCESSION BU619082
VERSION BU619082.1 GI:23285297
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Martin

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this CDNA
sequence: 1-32, >AT_rich#low complexity (matched complement)
298-324, >AT_rich#low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source Location/Qualifiers
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FH1-bfm-f-06-0-UI"
/tissue_type="Cell Line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_FH1"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia
I; NCI CGAP_FH1 is a normalized cDNA library obtained from
a cell line derived from grade I chondrosarcoma tissue.
The library was constructed and normalized according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGAATCCGGC. The cell line
was provided by Dr. James Martin from the University of
Iowa.
TAG_LIB=UI-H-FH1
TAG_TISSUE=Human Chondrosarcoma Cell Line CS8 - Grade 1
Chondrosarcoma
TAG_SEQ=AGAATCCGGC"

BASE COUNT 253 a 99 c 89 g 271 t
ORIGIN

Query Match 91.5%; Score 653.4; DB 13; Length 712;
Best Local Similarity 97.0%; Pred. No. 1.4e-67;
Matches 674; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 13 TTATTTTAAATTTCTATTAACATCTCTCAAGCATTATTTATCCATATCTCACTG 72
Db 12 TTATTTTAAATTTCTATTAACATCTCTCAAGCATTATTTATCCATATCTCACTG 71
QY 73 AATTTTAAAGAAATACATAGTATAGAAAACCTAGAAAAGATTAATGAGATAATT 132
Db 72 AATTTTAAAGAAATACATAGTATAGAAAACCTAGAAAAGATTAATGAGATAATT 131
QY 133 AAACCTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATTAATTGAAATGA 192
Db 132 AAACCTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATTAATTGAAATGA 191
QY 193 GATTATATTTGAAAACTGCATCTGAAAGCAACTTATTGTTCAATTAATNCTTAATGAT 252
Db 192 GATTATATTTGAAAACTGCATCTGAAAGCAACTTATTGTTCAATTAATNCTTAATGAT 251
QY 253 GGTGTTTATGACTAATACATGATTTTCAAGAGAAAACCATGTTAAAAATATTTT 312
Db 252 GGTGTTTATGACTAATACATGATTTTCAAGAGAAAACCATGTTAAAAATATTTT 311
QY 313 ATTTTAAAAATAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGAATTTGGAA 372
Db 312 ATTTTAAAAATAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGAATTTGGAA 371
QY 373 NAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTAATCTCACTAATT 432
Db 372 NAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTAATCTCACTAATT 431

Db 372 GAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTAAATCTCACTAATT 431
QY 433 TTANAACTATGAGAAATTTGATTAATGACATGACAGACACACTAATTAATGAGCCAG 492
Db 432 TTAAGAACTATGAGAAATTTGATTAATGACATGACAGAGTGCACACACTAATTAATGAGCCAG 491
QY 493 CTGTTGGCATTTGTTTCTTACTTACTAGTCTCTCCAAAGGAAAACTTTAAATGAAATCTTC 552
Db 492 CTGTTGGCATTTGTTTCTTACTTACTAGTCTCTCCAAAGGAAAACTTTAAATGAAATCTTC 550
QY 553 AGCAGAAATATCTTAAATATATCTTTGTAAGCAAAACAAAGCTTTTGTTCATAGT 612
Db 551 AGCAGAAATATCTTAAATATATCTTTGTAAGCAAAACAAAGCTTTTGTTCATAGT 610
QY 613 TCTTTGGCATTTTACTGTTCTCCTAATTTTATCTGAAACCTCAATTTTACCCGACCATAA 672
Db 611 TCTTTGGCATTTTACTGTTCTCCTAATTTTATCTGAAACCTCAATTTTACCCGACCATAA 670
QY 673 TTACCATATTAATCTTTGTTTNGACACAGTTGTTGC 707
Db 671 TTACCATATTAATCTTTGTTTNGACACAGTTGTTGC 705

RESULT 15
BQ009270
LOCUS
DEFINITION
UI-H-ED1-axx-n-02-0-UI.s1 NCI CGAP_ED1 Homo sapiens CDNA clone
IMAGE:5835073 3', mRNA sequence.
ACCESSION
BQ009270
VERSION
BQ009270.1 GI:19734171
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabds-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende.
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/INL at: http://image.llnl.gov
The following repetitive elements were found in this CDNA
sequence: 316-342, >AT_rich#low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5835073"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library
containing the following tissue(s): Chondrosarcoma cell
line CS5. The library was constructed according to Bonaldi
, Lennon and Soares, Genome Research, 6:791-806, 1996.
First strand cDNA synthesis was primed with an oligo-dT
primer containing a Not I site. Double stranded cDNA was
ligated to an EcoR I adaptor, digested with Not I, and
cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is GCTCAAGGCT.

TAG_LIB=UI-H-ED1
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"

BASE COUNT 266 a 105 c 96 g 281 t 2 others
ORIGIN

Query Match 91.2%; Score 651.2; DB 12; Length 750;
Best Local Similarity 97.1%; Pred. NO. 2.5e-67;
Matches 692; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

QY	2	GCCAGAAAAAGTTATTTTAAATTTCTATTAACATCTTCTCAAGCATTTATTCCT	61
Db	19	GCCAGAAAAAGTTATTTTAAATTTCTATTAACATCTTCTCAAGCATTTATTCCT	78
QY	62	ATATCTCAGTGAATTTTAAGAATAACATTAGATTAGAAAACTAGAAAAAAGATAA	121
Db	79	ATATCTCAGTGAATTTTAAGAATAACATTAGATTAGAAAACTAGAAAAAAGATAA	138
QY	122	TGCAGATATTAACCTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATA	181
Db	139	TGCAGATATTAACCTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATA	198
QY	182	AATTGAAATGAGATTATTAATTTGAAAACTGCACTGAAAGCAAACTTTATGTTCAATTA	241
Db	199	AATTGAAATGAGATTATTAATTTGAAAACTGCACTGAAAGCAAACTTTATGTTCAATTA	258
QY	242	TNCTTAATGATGGTGTTTTATGACTAATACATGATTTTCAAGAAAGAAACCATGTTA	301
Db	259	TTCTTAATGATGGTGTTTTATGACTAATACATGATTTTCAATTAAGAAACCATGTTA	318
QY	302	AAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATT	361
Db	319	AAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATT	378
QY	362	TGATTTGGGAANAAAACTACTGTTCTGATAGCAATGCAAAATTTTAGATTTTAA	421
Db	379	TGATTTGGGAANAAAACTACTGTTCTGATAGCAATGCAAAATTTTAGATTTTAA	438
QY	422	TCTCACTAATTTTAANAACATATTGAAATTTGATGACATGAAGTGACAACACTAA	481
Db	439	TCTCACTAATTTTAAGAACATATTGAAATTTGATGACATGAAGTGACAACACTAA	498
QY	482	TTACTGGCCAGCTGTGGCATTTGTTCTTAAGTTCTCCCAAGGAAAACTCTTAA	541
Db	499	TTACTGGCCAGCTGTGGCATTTGTTCTTAAGTTCTCCCAAGGAAAACTCTTAA	557
QY	542	ATTGAATCTTCAGCAGAAATATCCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT	601
Db	558	ACTGAATCTTCAGCAGAAATATCCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT	617
QY	602	GTTTACATAGTCTTTGGGATTTTACCTGTTCTAATTTTATTTCTGAACCTCAA-TTTTAC	660
Db	618	GTTTACATAGTCTTT-GGATTTTACCTGTTCTAATTTTANTCTGAACCTCAATTTTAC	676
QY	661	CCGAGACCATAATTACCATATTACTTTGTTNGACAGTTGTTGCAATTG 713	
Db	677	CCGAGACCATAATTACCATATTACTTTGTTNGACAGTTGTTGCAATTG 729	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:01:24 ; Search time 234.619 Seconds
(without alignments)
8215.019 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 714
Sequence: 1 cgccagaaaagctatttta.....cacagtgttgcacatcca 714

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03: *
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
--------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

c	9	126.2	17.7	397	25	ABx40797	Bovine EST associa
c	10	107.4	15.0	263	15	AAQ76489	Human genome fragm
c	11	77.8	10.9	8056	25	ABZ10246	Haematopoietic cel
c	12	76.8	10.8	397	25	ABx47851	Bovine EST associa
c	13	75.2	10.5	449	25	ABx48095	Bovine EST associa
c	14	73.4	10.3	2152	24	AB199762	Mouse ischaemic co
c	15	71	9.9	8056	25	ABZ10246	Haematopoietic cel
c	16	70.2	9.8	8056	25	ABZ10100	Haematopoietic cel
c	17	67.6	9.5	883	22	AA15210	Human breast cance
c	18	67.6	9.5	50000	24	ABL56201	Human genome fragm
c	19	67.4	9.4	4985	24	ABQ75107	Anopheles gambiae
c	20	66.4	9.3	5852	12	AAQ11710	Dicystostelium plas
c	21	65.2	9.1	7676	24	ABL70409	Chemically treated
c	22	65.2	9.1	7676	24	ABL34598	Human metastasis a
c	23	64.6	9.0	6106	22	AA546429	Tumour suppressor
c	24	64.6	9.0	6106	24	ABK40031	Human chemically p
c	25	64.6	9.0	6106	24	ABL33472	Human immune syste
c	26	64.6	9.0	13131	24	ABL92249	Chemically treated
c	27	64.2	9.0	8056	25	ABZ10100	Haematopoietic cel
c	28	63.6	8.9	6944	24	ABK34026	Human DNA for stag
c	29	62.8	8.8	1501	25	ABZ10188	Haematopoietic cel
c	30	62.8	8.8	50000	24	ABL56201	Human genome fragm
c	31	62.2	8.7	291	25	ABx48529	AMEV genome fragm
c	32	62.2	8.7	10048	24	ABL70313	Bovine EST associa
c	33	62.2	8.7	10048	24	AAS61251	Chemically treated
c	34	61.8	8.7	5962	24	ABL33287	Human gene regulat
c	35	61.4	8.6	17721	24	ABL33729	Human immune syste
c	36	61.2	8.6	20674	21	AAC58017	Archidonic acid m
c	37	60.8	8.5	424	25	ABx46053	Bovine EST associa
c	38	60.6	8.5	10467	24	ABK28453	DNA transcription
c	39	59.6	8.3	6169	22	AA546370	Tumour suppressor
c	40	59.6	8.3	6169	24	ABN80097	Human chemically m
c	41	59.6	8.3	6912	24	ABK28371	DNA transcription
c	42	59.6	8.3	7631	24	ABL32861	Human immune syste
c	43	59.4	8.3	12138	24	ABL33943	Human immune syste
c	44	59.4	8.3	12138	24	ABK28336	DNA transcription
c	45	59.2	8.3	5815	24	ABK40024	Human chemically p

ALIGNMENTS

RESULT 1		
ID	AAx40094	standard; DNA; 714 BP.
XX	AAx40094	
AC	AAx40094;	
XX	02-JUL-1999	(first entry)
DT		
XX		
DE		Gastric cancer associated gene.
XX		
KW		Cancer associated antigen; diagnosis; research; treatment; human;
KW		breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW		prostate cancer; ss.
XX		
OS		Homo sapiens.
XX		
PN	WO9904265-A2.	
XX		
PD	28-JAN-1999.	
XX		
PF	15-JUL-1998;	98WO-US14679.
XX		
XX	22-JUN-1998;	98US-0102322.
PR	17-JUL-1997;	97US-0896164.
PR	10-OCT-1997;	97US-0061599.
PR	10-OCT-1997;	97US-0061765.
PR	10-OCT-1997;	97US-0948705.
PR	11-OCT-1997;	97GB-0021697.
XX		
PA	(LUDW-)	LUDWIG INST CANCER RES.
XX		

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ,
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
PS Claim 67; Page 695; 787pp; English.
XX
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
SQ Sequence 714 BP; 258 A; 100 C; 92 G; 260 T; 4 other;

Query Match 99.4%; Score 710; DB 20; Length 714;
Best Local Similarity 100.0%; Pred. No. 1.8e-113;
Matches 714; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCAGAAAAAGTATTTTAATTTCTATTAACATCTTCTCAAGCATTTTATCC 60
DB 1 CGCCAGAAAAAGTATTTTAATTTCTATTAACATCTTCTCAAGCATTTTATCC 60
QY 61 TATATCTCACTGAATTTTAGAAATAACATTAGTATTAGAAAACTAGAAAAAGATTA 120
DB 61 TATATCTCACTGAATTTTAGAAATAACATTAGTATTAGAAAACTAGAAAAAGATTA 120
QY 121 ATGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGACGTTAT 180
DB 121 ATGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGACGTTAT 180
QY 181 AAATGGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATT 240
DB 181 AAATGGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATT 240
QY 241 ATNCTTAATGATGTTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCATGTT 300
DB 241 ATNCTTAATGATGTTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCATGTT 300
QY 301 AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATT 360
DB 301 AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATT 360
QY 361 TTGATTGGGAANAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTGA 420
DB 361 TTGATTGGGAANAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTGA 420
QY 421 ATCTCACTAATTTTAAANAATACTATGAGAAATGATTAATGACATGAAGTGCACAACTA 480
DB 421 ATCTCACTAATTTTAAANAATACTATGAGAAATGATTAATGACATGAAGTGCACAACTA 480
QY 481 ATTACTGGCCAGCTGTGGCATGTGTTCTTACTAGTCTCTCCCAAGGAAAACTCTTA 540
DB 481 ATTACTGGCCAGCTGTGGCATGTGTTCTTACTAGTCTCTCCCAAGGAAAACTCTTA 540
QY 541 AATTGAATCTTCAGCAGAAATAATCCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT 600
DB 541 AATTGAATCTTCAGCAGAAATAATCCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT 600

QY 601 TGTTCATAGTCTTTGGGATTGTTCTTAATTTATTTATCTGAACCTCAATTTTAC 660
DB 601 TGTTCATAGTCTTTGGGATTGTTCTTAATTTATTTATCTGAACCTCAATTTTAC 660
QY 661 CCCAGACCATAATTACCATATTAATCTTTGTTTGACACAGTTGTTGCCAATTCA 714
DB 661 CCCAGACCATAATTACCATATTAATCTTTGTTTGACACAGTTGTTGCCAATTCA 714

RESULT 2
AAH57501/c
ID AAH57501 standard; cDNA; 2663 BP.
XX
AC AAH57501;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human liver cell specific cDNA sequence SEQ ID NO:341.
XX
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS Homo sapiens.
XX
PN WO200132927-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000WO-US30396.
XX
PR 04-NOV-1999; 99US-0163508.
XX
PA (INCY-) INCYTE GENOMICS INC.
PI Sornasse T, Selhammer J, Watson GA;
XX
DR WPI; 2001-291057/30.
XX
PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
XX
PS Claim 1; Page 258-259; 327pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.
XX
SQ Sequence 2663 BP; 807 A; 472 C; 599 G; 785 T; 0 other;

Query Match 94.5%; Score 674.4; DB 22; Length 2663;
Best Local Similarity 98.3%; Pred. No. 2.3e-107;
Matches 700; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 2 GCCAGAAAAAGTATTATTATTTCTAATTAACATCTCTCTCAAGCATTTATCTCT 61
DB 2652 GCCAGAAAAAGTATTATTATTTCTAATTAACATCTCTCTCAAGCATTTATCTCT 2593

QY 62 ATATCTCACTGAATTTAAGAATAACATAGTATTAGAAAACTAGAAAAAGATATA 121
|||||
DB 2592 ATATCTCACTGAATTTAAGAATAACATAGTATTAGAAAACTAGAAAAAGATATA 2533
QY 122 TGCAGATAATTAACCTACATGAAAAAGAAATTAACAAGAGAGCTGAGACGTATATA 181
|||||
DB 2532 TGCAGATAATTAACCTACATGAAAAAGAAATTAACAAGAGAGCTGAGACGTATATA 2473
QY 182 AATTGAATGAGATTATAATTGAAAACCTGCTGAAAGCAAACTTTATTGTTCAATTA 241
|||||
DB 2472 AATTGAATGAGATTATAATTGAAAACCTGCTGAAAGCAAACTTTATTGTTCAATTA 2413
QY 242 TNCCTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAGGAAACCCATGTTA 301
|||||
DB 2412 TTCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAGGAAACCCATGTTA 2353
QY 302 AAAATATTTTATTTTAAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 361
|||||
DB 2352 AAAATATTTTATTTTAAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 2293
QY 362 TGATTTGGGAANAATACTGTTCTGATAGCAATGAATGCAAAATTTTATGATTTTAA 421
|||||
DB 2292 TGATTTGGGAANAATACTGTTCTGATAGCAATGAATGCAAAATTTTATGATTTTAA 2233
QY 422 TCTCACTAATTTTANANAATACTGAGAAATGATTAATGACATGAGTGACACACTAA 481
|||||
DB 2232 TCTCACTAATTTTANANAATACTGAGAAATGATTAATGACATGAGTGACACACTAA 2173
QY 482 TTACTGGCCAGCTGTGGCATTGTGTTCTTACTAGTTCTCCCAAGGAAAACTCTTAA 541
|||||
DB 2172 TTACTGGCCAGCTGTGGCATTGTGTTCTTACTAGTTCTCCCAAGGAAAACTCTTAA 2114
QY 542 ATTGAATCTTCAGACAGATAATCCTTAATATATCTTGTAAAGCAAAACAAAGCTTTT 601
|||||
DB 2113 ACTGAATCTTCAGACAGATAATCCTTAATATATCTTGTAAAGCAAAACAAAGCTTTT 2054
QY 602 GTTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACCTCAATTTACC 661
|||||
DB 2053 GTTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAAACCTCAATTTACC 1995
QY 662 CCAGACCATTAATTACCATTTAATCTTTGTCAGACAGTTGTTGCCAATTC 713
|||||
DB 1994 CCAGACCATTAATTACCATTTAATCTTTGTCAGACAGTTGTCAGACAGTTGTCAGACAGTTGTC 1943
RESULT 3
AAK40095
ID AAK40095 standard; DNA; 687 BP.
XX
AC AAK40095;
XX
DT 02-JUL-1999 (first entry)
XX
DE Gastric cancer associated gene.
KW Cancer associated antigen; diagnosis; research; treatment; human;
KM breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
XX prostate cancer; es.
OS Homo sapiens.
XX
PN WO9904265-A2.
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98WO-US14679.
XX
PR 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.

XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfeundschnuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
PS Claim 67; Page 695; 787pp; English.
XX
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
SQ Sequence 687 BP; 242 A; 98 C; 85 G; 258 T; 4 other;
Query Match 85.7%; Score 611.8; DB 20; Length 687;
Best Local Similarity 97.1%; Pred. No. 1.3e-96;
Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;
QY 15 ATTTAATTTTCTATTAAACATTTCTCAAAAGCATTTATTTATCTCAATCTCACTGAA 74
|||||
DB 1 ATTTAATTTTCTATTAAACATTTCTCAAAAGCATTTATTTATCTCAATCTCACTGAA 60
QY 75 TTTAAGAAATAACATTAGTATTAGAAAACTAGAAAAAGATAAATGAGATAATTAA 134
|||||
DB 61 TTTAANAATAACATTAGTATTAGAAAACTAGAAAAAGATNAATGAGATAATTAA 120
QY 135 ACTTACATGAAAAAGGAAATTTATAACAAGAGAGCTGAGAACGTTATTAATTGAATGAGA 194
|||||
DB 121 ACTTACATGAAAAAGGAAATTTATAACAAGAGAGCTGAGAACGTTATTAATTGAATGAGA 180
QY 195 TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG 254
|||||
DB 181 TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG 240
QY 255 TGTTTATGACTAATAACACTGATTTTTCAGAGAAAGAAACCATGTTAAATAATTTTAT 314
|||||
DB 241 TGTTTATGACTAATAACACTGATTTTTCAGAGAAAGAAACCATGTTAAATAATTTTAT 300
QY 315 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTTGGGAANA 374
|||||
DB 301 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTTGGGAANA 360
QY 375 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAACTCACTAATTTT 434
|||||
DB 361 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAACTCACTAATTTT 420
QY 435 AANAATATGAGAAATGATTAATGACATGAAGTGACACACACTAATTAATGAGCCAGCT 494
|||||
DB 421 AAGAATATGAGAAATGATTAATGACATGAAGTGACACACACTAATTAATGAGCCAGCT 480
QY 495 GTTGCAATGTTGTTCTTACTAGTTCTCCCAAGGAAAACTTTAAATGAAATCTTCAG 554
|||||
DB 481 GTTGCAATGTTGTTCTTACTAGTTCTCCCAAGGAAAACTTTAAATGAAATCTTCAG 539
QY 555 CAGATAATCTTTAATAATATACTTTGTAAGCAAAACAAAGCTTTTGTGTTTACATAGTTC 614

Db 540 CNGAATAA-CCTTAATATACTTTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTC 597
QY 615 TTGGGATTTTACTGTCCCTAATTTTATCTGAAACTCAATTTTACCACGACCATTAAT 674
Db 598 TTT-GGATTTTACTGTCTTAATTTTATCTGAAACTCCATTTTCCCAGACCATTAAT 656
QY 675 ACCATA-TTAACTTGTGTTGACAGTTGTT 704
Db 657 ACCCTATTTAACCTTGTATGACAGTTGTT 687
RESULT 4
ABV87267
ID ABV87267 standard; cDNA; 486 BP.
XX
AC ABV87267;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human colon cancer related cDNA SEQ ID NO 578.
XX
KM Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200258534-A2.
XX
PD 01-AUG-2002.
XX
PF 19-NOV-2001; 2001WO-US43704.
XX
PR 20-NOV-2000; 2000US-252222P.
PR 06-FEB-2001; 2001US-267011P.
PR 28-MAR-2001; 2001US-279670P.
PR 10-JUL-2001; 2001US-304037P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
XX
DR WPI; 2002-608400/65.
XX
PT New isolated tumor colon polynucleotide and polypeptide, useful for the
PT diagnosis, prevention and/or treatment of cancer, in particular colon
PT cancer -
XX
PS Claim 1; SEQ ID NO 578; 266bp + Sequence Listing; English.
XX
CC The invention relates to a human colon tumour expressed polynucleotide
CC (I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of
CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
CC complements of (i); (iii) at least 20 contiguous residues of (i); (iv)
CC sequences that hybridize to (i), under moderately stringent conditions;
CC (v) sequences having at least 75% or 90% identity to (i); or (vi)
CC degenerate variants of (i). The compositions and methods of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC cancer, particularly colon cancer. (I) can be used in gene therapy and
CC (I) and (II) are useful in pharmaceutical compositions such as vaccines.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 486 BP; 192 A; 55 C; 64 G; 175 T; 0 other;
Query Match 67.0%; Score 478.2; DB 24; Length 486;
Best Local Similarity 98.8%; Pred.No.1.2e-73;
Matches 480; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CCAGAAAAAGTTATTTAATTTCTATTAACATCTCTCAAGCATTTATTCCTA 62
XX
PT 1 CCAGAAAAAGTTATTTAATTTCTATTAATTAATTTCTCTCAAGCATTTATTCCTA 60
Db

QY 63 TATCTCACTGAATTTTAAGAAATACATTAAGTATTAGAAAAACTAGAAAAAGATAAT 122
Db 61 TATCTCACTGAATTTTAAGAAATACATTAAGTATTAGAAAAACTAGAAAAAGATAAT 120
QY 123 GCAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAAGACTGAGAACGTTATAA 182
Db 121 GCAGATAATTAACTTACATGAAAAAGAAAAATTATAACAAAGACTGAGAACGTTATAA 180
QY 183 ATTGAATGAGATTATAATTGAAAACTGCACTGAAAGCAAACTTTAATTGTTCAATTAT 242
Db 181 ATTGAATGAGATTATAATTGAAAACTGCACTGAAAGCAAACTTTAATTGTTCAATTAT 240
QY 243 NCTTAATGATGCTGTTTATGACTAATACTGATTTTCAAGAGAAACCATGTAA 302
Db 241 TCTTAATGATGCTGTTTATGACTAATACTGATTTTCAATTAAGAAACCATGTAA 300
QY 303 AAATATTTTATTTTAAATAAGCCCTGCTGTTCAAGCTCTGATCATATTTCTTTATTTT 362
Db 301 AAATATTTTATTTTAAATAAGCCCTGCTGTTCAAGCTCTGATCATATTTCTTTATTTT 360
QY 363 GATTTGGGAANAATACTGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAAT 422
Db 361 GATTTGGGAANAATACTGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAAT 420
QY 423 CTCACATAATTTTAAANAACATATTGAGAAATGATTAATGACATGAAGTCAACACTAAT 482
Db 421 CTCACATAATTTTAAAGAATATTGAGAAATGATTAATGACATGAAGTCAACACTAAT 480
QY 483 TACTGG 488
Db 481 TACTGG 486

RESULT 5
AAX40097
ID AAX40097 standard; DNA; 843 BP.
XX
AC AAX40097;
XX
DT 02-JUL-1999 (first entry)
XX
DE Gastric cancer associated gene.
XX
KM Cancer associated antigen; diagnosis; research; treatment; human;
KM breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9904265-A2.
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98WO-US14679.
XX
PR 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Gout I, Guire A, O'Hare M, Obata Y, Old LJ;
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products

[illegible]

RESULT 7
AAH57293/c
ID AAH57293 standard; cDNA; 262 BP.

AC AAH57293;

DT 10-SEP-2001 (first entry)

DE Human liver specific cDNA sequence SEQ ID NO:133.

KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

OS Homo sapiens.

PN WO200132927-A2.

PD 10-MAY-2001.

PF 02-NOV-2000; 2000WO-US30396.

PR 04-NOV-1999; 99US-0163508.

PA (INCY-) INCYTE GENOMICS INC.

PI Sornasse T, Seilhamer JJ, Watson GA;

DR WPI; 2001-291057/30.

PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -

PS Claim 1; Page 117; 327pp; English.

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide sequences (I). (I) can have cytostatic, immunomodulatory and neuroprotective activities, and can be used in gene therapy. (I) and proteins (II) encoded by them are used in high throughput screening assays to select DNA molecules, RNA molecules, peptide nucleic acids, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents. Expression of (I) in a sample indicates the differentiation of embryonic stem cells into a tissue selected from brain, heart, kidney, liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used to produce an expression profile that defines a metabolic or developmental process, treatment, condition, disease or disorder. The gene profile can be used for diagnosis, prognosis or monitoring of treatments and for investigating a predisposition to a disorder where the gene is associated with a cancer, immunopathology or neuropathology.

SQ Sequence 262 BP; 103 A; 38 C; 34 G; 87 T; 0 other;

Query Match	36.2%	Score 258.4;	DB 22;	Length 262;
Best Local Similarity	98.9%	Pred. No. 7.1e-36;		
Matches 259; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	257	TTTTATGACTAATACACTGATTTTTCAGAAGGAAACCCATGTTAAAAATATTTTATT	316
Db	262	TTTTATGACTAATACACTGATTTTTCATAAGGAAACCCATGTTAAAAATATTTTATT	203
QY	317	TAAAAATAAGCCTGTGTTCAAGCTGATCATATTTCTTTATTTTGATTGGGAANAAA	376
Db	202	TAAAAATAAGCCTGTGTTCAAGCTGATCATATTTCTTTATTTTGATTGGGAAGAAA	143
QY	377	ATACTGTTTCTGATAGCATGAAATGCAAAATTTTGTAGATTTTAACTCACTAATTTAA	436
Db	142	ATACTGTTTCTGATAGCATGAAATGCAAAATTTTGTAGATTTTAACTCACTAATTTAA	83
QY	437	NAACTATTGAGAAATTGATTAAATGACATGAAGTGCACAACAATAATTACTGGCCAGCTGT	496
Db	82	GAACATAATTGAGAAATTGATTAAATGACATGAAGTGCACAACAATAATTACTGGCCAGCTGT	23
QY	497	TGGCATTTGTTTCTTACTTAG	518
Db	22	TGGCATTTGTTTCTTACTTAG	1

RESULT 8
ABX83093/c
ID ABX83093 standard; cDNA; 173 BP.

AC ABX83093;

DT 24-APR-2003 (first entry)

Corn ear-derived polynucleotide (cpd) #1553.

KM Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022;
KM SATMON023; structural gene; functional gene; regulatory gene;
KM corn ear-specific profile; gene transcription; gene expression;
KM hybrid plant; desirable trait expression; plant breeding program;
KM inheritance; desired characteristic; growth; development;
KM disease resistance; environmental adaptability; quality; yield;
KM multigene trait; plant; gene; ss.

OS Zea mays.

PN US6476212-B1.

PD 05-NOV-2002.

PF 14-MAY-1999; 99US-0313294.

PR 26-MAY-1998; 98US-086722P.

PA (INCY-) INCYTE GENOMICS INC.

PI Lalgudi RV, Ito LY, Sherman BK;

DR WPI; 2003-208840/20.

PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with
PT growth, development -

PS Examples; SEQ ID No 1553; 390pp; English.

The present invention relates to the isolation of corn ear-derived polynucleotides (cdps) from the corn (*Zea mays*) cDNA libraries SATMON022 and SATMON023. Some of the cpds uniquely identify structural, functional, and regulatory genes of corn ear. The polynucleotides sequences are useful for detecting cpds in a sample, for producing a corn ear-specific profile of gene transcription, for detecting altered gene expression in inbred or hybrid plants, and for screening several molecules for specific binding to the polynucleotide. The cpds are useful to identify, isolate, or extend identical or related corn-ear nucleic acid sequences from DNA libraries, and in nucleic acid amplification or hybridisation techniques to follow the

CC expression of desirable traits through plant breeding programs.
CC Preferably, the cdds are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality, and yield of corn. The cdds are also useful as molecular
CC markers for studying inheritance and multigene traits in a plant
CC breeding program. The cdds are useful for producing purified corn-ear
CC polypeptides by recombinant techniques. They are also useful in
CC diagnostic assays to detect or confirm conditions or diseases
CC associated with abnormal levels of cdp expression. ABX81541-ABX89140
CC represent corn ear-derived polynucleotides (cpds) of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/patseq/entry.html.
XX
SQ Sequence 173 BP; 54 A; 20 C; 19 G; 80 T; 0 other;
Query Match 23.2%; Score 165.6; DB 25; Length 173;
Best Local Similarity 97.7%; Pred. No. 6.2e-20;
Matches 168; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 8 AAAAGTTATTTAATTTCTATTAAACATCTCTCAAGCATTTATTTCTATATCT 67
DB 172 AAAAAGTTATTTAATTTCTATTAAACATCTCTCAAGCATTTATTTCTATATCT 113
QY 68 CACTGAATTTTGAATAATACATTAGTATTAGAAAACTAGAAAAAAGATAATGCAGA 127
DB 112 CACTGAATTTTGAATAATACATTAGTATTAGAAAACTAGAAAAAAGATAATGCAGA 53
QY 128 TAATTAACCTTACATGAAAAAGAAATTTATACAAAGACTGAGACGTTA 179
DB 52 TAATTAACCTTACATGAAAAAGAAATTTATACAAAGACTGAGACGTTA 1
RESULT 9
ABX40797 standard; cDNA; 397 BP.
XX
AC ABX40797;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #5962.
XX
KW Bovine; 88; EST; expressed sequence tag; lactation; LMPD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
PS Claim 2; SEQ ID No 5962; 245bp; English.

XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived
CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMPD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMPD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.
XX
SQ Sequence 397 BP; 144 A; 75 C; 45 G; 133 T; 0 other;
Query Match 17.7%; Score 126.2; DB 25; Length 397;
Best Local Similarity 73.6%; Pred. No. 3.7e-13;
Matches 204; Conservative 0; Mismatches 60; Indels 13; Gaps 3;
QY 435 AANAACTATGAGAAATTGATTAATGACATGAGTGACACACACTAATTAAGCCAGCT 494
DB 14 AATTATCACTGAGAAATATATTAAAGGTCAATTAATGCACACACTTAATTAAGTCAAGCT 73
QY 495 GTTGCAATGTTGTTCTTACTTACTTCTCCCAAGGAAAGCTTTAAATTGAATCTTCAG 554
DB 74 AATATGATAT-TCTGTTCTTCATCTCCCAAGAAAC-----TTGAAATTTTCGG 124
QY 555 CAGATAATCTTAAATATATCTTTGTAGCAAAACAAAGCTTTTGTTCATAGTTC 614
DB 125 CAGATAATCTTCAATATATACATTATTAGCAAAATGAGAGCTTCTGTTT---ACATACT 180
QY 615 TTGGGATTTTACTGTTCTCTAATTTTATCTGAAACTCAATTTTACCCAGACCATTAAT 674
DB 181 TTTGTATTTTGGTATTTCTAATCTTATCTTAAAACTCAATTTTACCCCAACCATGATT 240
QY 675 ACCATATTACTTTGTNTGACACAGTTGTTGCCAAT 711
DB 241 ACCATATTACTTTGTATATGACACAGTTGTTGCAACT 277
RESULT 10
AAQ76489/c
ID AAQ76489 standard; DNA; 263 BP.
XX
AC AAQ76489;
XX
DT 25-MAR-2003 (updated)
DT 23-SEP-1994 (first entry)
XX
DE Human genome fragment.
XX
KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
KW detection; homology; human; adrenal tissue; ds.
XX
OS Homo sapiens.
XX
PN WO9401548-A2.

XX 20-JAN-1994.
PD
XX
XX 13-JUL-1993; 93WO-GB01467.
PF
XX
XX 13-JUL-1992; 92GB-0014857.
PR
XX
XX (MEDI-) MEDICAL RES COUNCIL.
PA
XX Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;
PI Sibson DR, Starkey M;
XX WPI; 1994-035056/04.
DR
XX
XX New nucleic acid fragment encoding gene products - can be used
PT for genetic analysis and mapping
XX
XX Claim 1; Page 64; 616pp; English.
PS
XX Human nucleic acid fragments, isolated from brain adrenal tissue,
CC the placenta or bone marrow comprise any of: (A) a sequence
CC selected from (AAQ6401-Q7613), (B) an allelic variation of a
CC sequence as described in (A), or (C) a sequence complementary
CC to (A) or (B).
CC (Updated on 25-MAR-2003 to correct PN field.)
CC
XX Sequence 263 BP; 97 A; 38 C; 45 G; 80 T; 3 other;
SQ
Query Match 15.0%; Score 107.4; DB 15; Length 263;
Best Local Similarity 82.1%; Pred. No. 6.3e-10;
Matches 207; Conservative 0; Mismatches 34; Indels 11; Gaps 7;
QY 464 TGAAGTGACACACACTAATTAAGCCAGCTGTGGCATTTGTTCTTACTTACTTCTC 523
Db 263 TGGAGTACACAGACACTGATTAAGCCAGCTGTGGCATTTGTTCTTACTTACTTCTC 205
QY 524 CCAAGGAAAACTCTTAATTTGAATCTTCAGCAGATAATCCTTAATATATCTTTGTAAG 583
Db 204 ACGA-GGAAACTCTTAATCTGAATCTT--AGCAGAAATCCTTAATATATCTTTGTAAG- 150
QY 584 CAAACCAAAAGCTTTTGTGTTTACATAGTTCTTGGATTTTACTGTTCTTAATTTATT 643
Db 149 CAAACCAAAAGCTTTTGTGTTTACATAGTTCTTGGATTTTACTGTTCTTAATTTATT 91
QY 644 CTGAAACTCAATTTTAC-CCAGACCATTAATTAATTAATTTGTTTGGACAGTTG 702
Db 90 CTGAAACTCGATTTTACNNCCAAACCATTAATTAATTTGTTTGGACAGTTG 34
QY 703 TTGGCAATTCA 714
Db 33 TATGCAATTGCA 22
RESULT 11
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.
XX
XX AC ABZ10246;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
XX Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX

PF 26-MAR-2002; 2002WO-EP03401.
XX
XX 26-MAR-2001; 2001US-278333P.
PR
XX
XX (EPIC-) EPIGENOMICS AG.
PA
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Diepenbrock C, Adorjan P, Grabs G, Ilesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
PI Pelet C, Schwobe I, Ziebarth H;
XX WPI; 2003-018942/01.
DR
XX
XX Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent
PT that distinguishes between methylated and non-methylated CpG
PT dinucleotides -
XX
XX Claim 28; SEQ ID 386; 117pp; English.
PS
XX The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related
CC DNA sequences. The nucleotide sequences from the present invention can
CC also be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables
CC a highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients.
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 other;
Query Match 10.9%; Score 77.8; DB 25; Length 8056;
Best Local Similarity 47.3%; Pred. No. 7.9e-05;
Matches 338; Conservative 0; Mismatches 361; Indels 16; Gaps 3;
QY 7 AAAAAAGTTATTTAATTTCTATTAACATTTCTTCAAGCATTTTATCTATATC 66
Db 1368 AAAAAATTTATTTTATTAATAAATAATGATTTTATTTTATTTTATTTTATTTT 1427
QY 67 TCACTGAATTTAAGAATAACATTAATTTAGAAAACTAGAAAAAGATAA----- 120
Db 1428 TTTTAAAAAATAAATAATTTTATTAATAAATAATTAATAAATAATTAATTTAAT 1487
QY 121 ATGCAGATTAATTAACCTTACATGAAGAAAGAAATTAACAAAGAGCTGAGACGTTAT 180
Db 1488 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1547
QY 181 AAATTGAATGAGATTATTAATTTGAAAGCTGAAAGCAACTTATTTGTTCAATT 240
Db 1548 TTATTTTAAAAAATTAATTAATTTAATTAATTTTAAATTTAATTTAATTTAATTTAAT 1607
QY 241 ATNCTTAATGATGTTTATGACATAATACATGATTTTTCAGAGAAAGAACCATGTT 300
Db 1608 ATTTTAAAAAATGAAATTAATGAAATTAATTAATTAATTAATTAATTAATTAATG 1667
QY 301 AAAAA-----TATTTTATTTTAAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTT 356
Db 1668 ATAAAAATTTATTTTATTTAATAAATAATTAATTTAATTAATTAATTAATTAATTTAT 1727

Oy	357	TAATTTGATTTGGGAANNAATACTGTTTCTGATAGCATGAATGCAAAATTTTAGATT	416
Db	1728	AAATTTAATTATTTAAAAAATTAATAAAAAAATTTTAAATTAAAAAATTAATAATT	1787
Oy	417	TTTAATCTCACTAATTTTAAANAACTATTTGAGAAATTGATTAAAGACATGAAGTGACAAC	476
Db	1788	AAATTATGTTATTTATTTTAAATTAAATAATAAATAAATTTGTTTAAAAAATAA	1847
Oy	477	ACTAATTAATCTGGCCAGCTGTTGGCAATTGCTTTCTTACTTACTCTCCCAAGGAAAAT	536
Db	1848	AAAAAAAAAAAAAATTAATTTAAATTAATTAATTTTATTTTAAATAAATAAATAA---	1904
Oy	537	CTTAATTTGAATCTTCAGCAGAAATACTCTTAATATATCTTTGTAAGCAAAACAAAAGCT	596
Db	1905	---AAATTTTATAAAAAATAAATAAATAATTAATATATAAATAAATAAATAATTAATTT	1961
Oy	597	TTTTTGTTCACATAGTCTTTGGGATTTTACTGTTCTCTAATTTTAATCTGAAACTCAATT	656
Db	1962	TTTAAATAATAATATTAATATTAATTTTATTTTAAATTTAATAAATTTTAAATATATT	2021
Oy	657	TTACCCCAAGACCATTAATTACCAATACTTGTGTTNTGCACAGTTGTTGCCAAT	711
Db	2022	TAAAAATAAATAATATATTTTTAAATAATTTTATTTTAAATAAATTTTAAATAAT	2076
RESULT 12			
ABX47851/c			
ID	ABX47851	standard; cDNA; 397 BP.	
XX			
AC	ABX47851;		
XX			
DT	21-FEB-2003	(first entry)	
XX			
DE		Bovine EST associated with lactation/muscle/fat deposition #13016.	
XX			
KW		Bovine; ss; EST; expressed sequence tag; lactation; LMPD;	
KW		muscle deposition; fat deposition; genome mapping; gene identification;	
KW		gene analysis; cattle breeding.	
XX			
OS		Bos Taurus.	
XX			
PN		US2002137139-A1.	
XX			
PD		26-SEP-2002.	
XX			
PF		24-SEP-2001; 2001US-0960352.	
XX			
PR		12-JAN-1999; 99US-115707P.	
PR		11-JAN-2000; 2000US-0480902.	
XX			
PA		(BYAT/) BYATT J C.	
PA		(MATH/) MATHIALAGAN N.	
PA		(TAON/) TAO N.	
PA		(WARR/) WARREN W C.	
XX			
PI		Byatt JC, Mathialagan N, Tao N, Warren WC;	
XX			
DR		WPI; 2003-110599/10.	
XX			
PT		New nucleic acid associated with lactation, and muscle and fat	
PT		deposition, useful for genome mapping, gene identification and	
PT		analysis, cattle breeding, or for genetically improving cattle	
XX			
PS		Claim 2; SEQ ID No 13016; 245bp; English.	
XX			
CC		The invention relates to a purified nucleic acid molecule associated with	
CC		lactation or muscle and fat deposition (designated LMPD), derived	
CC		from cattle, and the LMPD nucleic acid can specifically hybridise to a	
CC		second nucleic acid molecule comprising any of 15112 nucleotide	
CC		sequences, appearing as ABX34836-ABX49947, or complements of them.	
CC		Also included are; (1) a transformed cell having a nucleic acid	
CC		comprising an LMPD nucleic acid linked to a promoter and a 3' non-	
CC		translated sequence that functions in the cell to cause termination of	

transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule.

The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMFD EST (expressed sequence tag) nucleic acids.

Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site: seqdata.uspto.gov/sequence.html?DocID=20020137139.

Query Match	10.8%;	Score 76.8;	DB 25;	Length 397;
Best Local Similarity	76.4%;	Pred. No. 0.00011;		
Matches 120; Conservative	0;	Mismatches 33;	Indels 4;	Gaps 2

QY		555 CAGATAATCCTTAAATATACCTTTGTGAAGCAAAACAAAAGCTTTTGTGTTACATAGTTC	614
Db	397	CAGAATAGTCTTCAAATATACATTATTAGCAAAATGAGAGC--TTCTGTTACATACTCA	3400
QY	615	TTTGGAATTTTACTGTCTCCTAATTTTAATCTGAAACTCAATTTTACCACCGACAATAATT	674
Db	339	T-T-GTATTCTGCTATTATTATACTATCTTTAAAACTCACTTTTACCCCAACCATGATTT	2822

QY	675	ACCATATTA	CTTGTG	TNTGC	AGTGTG	TGGCAAT	711
Db	281	ACCATATTA	ACTGTGT	ATGCA	CAGATG	TTAGCAACT	245

RESULT	13
ID	ABX48095/c
XX	ABX48095 standard; cDNA; 449 BP.
AC	ABX48095;
DT	21-FEB-2003 (first entry)
DE	Bovine EST associated with lactation/muscle/fat deposition #13260.
XX	
KW	. Bovine; BB; EST; expressed sequence tag; lactation; LMFJ;
KW	muscle deposition; fat deposition; genome mapping; gene identification;
KW	gene analysis; cattle breeding.
XX	
OS	Bos Taurus.
XX	
PN	US2002137139-A1.
XX	
PD	26-SEP-2002.
XX	
PF	24-SEP-2001; 2001US-0960352.
XX	
PR	12-JAN-1999; 99US-115707P.
PR	11-JAN-2000; 2000US-0480902.
XX	
PA	(BYAT/) BYATT J C.
PA	(MATH/) MATHIALAGAN N.
PA	(TAON/) TAO N.
PA	(WARR/) WARREN W C.
XX	
PI	Byatt JC, Mathialagan N, Tao N, Warren WC;
XX	
DR	WPI; 2003-110599/10.
XX	

PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle -
XX
PS Claim 2; SEQ ID No 13260; 245bp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived
CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 1512 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 1512 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMPD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 1512 bovine
CC LMPD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.
CC
SQ Sequence 449 BP; 148 A; 70 C; 89 G; 142 T; 0 other;
Query Match 10.5%; Score 75.2; DB 25; Length 449;
Best Local Similarity 85.6%; Pred. No. 0.00022;
Matches 83; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 615 TTGGGATTCTACTGTTCTTAATTTATTTCTGAAGCTCAATTTTACCCAGACCAATATT 674
Db 415 TTTGTATTTTGTATTCTTAATTTCTTAATTTCTGAAGCTCAATTTTACCCAGACCAATATT 356
QY 675 ACCATATTAACTTTGTTTTCACAGCTGTTGCCAAT 711
Db 355 ACCATATTAACTGTGTAATGCACAGTTGTTGCAACT 319
RESULT 14
ABI99762/c
ID ABI99762 standard; cDNA; 2152 BP.
XX
AC ABI99762;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:842.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UNINT-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR P-PSDB; ABB57301.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 2087-2091; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 2152 BP; 586 A; 455 C; 576 G; 535 T; 0 other;
Query Match 10.3%; Score 73.4; DB 24; Length 2152;
Best Local Similarity 66.7%; Pred. No. 0.00045;
Matches 180; Conservative 0; Mismatches 82; Indels 8; Gaps 5;

QY 420 AATCTCACTAATTTTAANAACCTATTGAGAAATGATTAATGACATGAGTGACAACACT 479
Db 2145 AATTTATTATTTAAGAACTATTAGTGAAGAACTGATGCAAGTAATTGAGTGAGGCACT 2086
QY 480 AATTACTGGCCAGCTGTTGGCATTGTGTTCTTAAGTTAGTTCTCCCAAGGAAACTCTT 539
Db 2085 AATTCTGTGTAACCT-ATGGAATCTGTTCTCTCTAGTCCCCCAAGAGAAA-TGTC 2028
QY 540 AATTGAATCTTCAGAGATAATATCCTTAATATACTTTGTAAGCAAAACAAGCTTTT 599
Db 2027 ACACCGAGTGTCCACACACAGTATCCTCAATATAC-TTCTAAGCAAAACAAGGCTTT- 1970
QY 600 TTGTTACATAGTCTTTGGGATTTTACTGTTCTTAATTTATTTCTGAAACTCAA-TTTT 658
Db 1969 ---GTTTACTCAGTCTTTCAGATTTTACTAGTCTGAAATTTTATTCTGAAACTCAAAGTTTA 1913
QY 659 ACCCCAGACCAATTAATCAATATTAACTTT 688
Db 1912 ACCCCAAACCAATCAATGACATAGTCAGCTT 1883
RESULT 15
ABZ10246/c
ID ABZ10246 standard; DNA; 8056 BP.
XX
AC ABZ10246;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 06:06:02 ; Search time 269.238 Seconds
(without alignments)
8737.270 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 714
Sequence: 1 cgccagaaaagtcatttca.....cacagctgttcgcaattca 714

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	710	99.4	714	9 US-09-835-992A-19	Sequence 19, Appl
2	672.8	94.2	1052	10 US-09-880-107-3296	Sequence 3296, Ap
3	611.8	85.7	687	9 US-09-835-992A-20	Sequence 20, Appl
4	482.4	67.6	506	10 US-09-796-692-7228	Sequence 7228, Ap
5	482.4	67.6	506	14 US-10-040-862-7228	Sequence 7228, Ap
6	478.2	67.0	486	10 US-09-998-598-578	Sequence 578, App
7	436.2	61.1	843	9 US-09-835-992A-22	Sequence 22, Appl
8	188	26.3	510	14 US-10-102-524-161	Sequence 161, App
9	188	26.3	510	14 US-10-102-524-188	Sequence 188, App
10	188	26.3	510	14 US-10-102-524-645	Sequence 645, App
11	126.2	17.7	397	10 US-09-960-352-5962	Sequence 5962, Ap
12	76.8	10.8	397	10 US-09-960-352-13016	Sequence 13016, A
13	75.2	10.5	449	10 US-09-960-352-13260	Sequence 13260, A
14	67.6	9.5	960	14 US-10-198-846-6381	Sequence 6381, Ap
15	67.4	9.4	4985	12 US-10-056-405-10	Sequence 10, Appl
16	67.4	9.4	4985	14 US-10-094-240-10	Sequence 10, Appl

17	65.2	9.1	7676	12 US-10-240-485-151	Sequence 151, App
18	65.2	9.1	3673778	12 US-10-312-841-1	Sequence 1, Appli
19	64.6	9.0	6106	12 US-10-311-455-1445	Sequence 1445, Ap
20	63.6	8.9	6944	14 US-10-172-086-111	Sequence 111, App
21	62.2	8.7	291	10 US-09-960-352-13694	Sequence 13694, A
22	62	8.7	3673778	12 US-10-312-841-2	Sequence 2, Appli
23	61.8	8.7	5962	12 US-10-311-455-1260	Sequence 1260, Ap
24	61.4	8.6	17721	12 US-10-311-455-1702	Sequence 1702, Ap
25	61.4	8.6	3673778	12 US-10-312-841-1	Sequence 1, Appli
26	60.8	8.5	424	10 US-09-960-352-11218	Sequence 11218, A
27	60.6	8.5	10467	12 US-10-240-453-327	Sequence 327, App
28	59.6	8.3	6912	12 US-10-240-453-245	Sequence 245, App
29	59.6	8.3	7631	12 US-10-311-455-834	Sequence 834, App
30	59.4	8.3	12138	12 US-10-311-455-1916	Sequence 1916, Ap
31	59.4	8.3	12138	12 US-10-240-453-210	Sequence 210, App
32	59.2	8.3	7046	12 US-10-311-455-2090	Sequence 2090, Ap
33	58.8	8.2	113515	12 US-10-311-455-2147	Sequence 2147, Ap
34	58.6	8.2	480	10 US-09-960-352-5301	Sequence 5301, Ap
35	58.4	8.2	10151	12 US-10-311-455-2405	Sequence 2405, Ap
36	58.4	8.2	16724	12 US-10-311-455-1064	Sequence 1064, Ap
37	58.4	8.2	16724	12 US-10-240-485-90	Sequence 90, Appl
38	58.4	8.2	113515	12 US-10-311-455-2148	Sequence 2148, Ap
39	58.2	8.2	640681	10 US-09-790-988-1	Sequence 1, Appli
40	58	8.1	3673778	12 US-10-312-841-2	Sequence 2, Appli
41	57.8	8.1	539	14 US-10-198-846-1369	Sequence 1369, Ap
42	57.8	8.1	6389	12 US-10-311-455-1933	Sequence 1933, Ap
43	57.6	8.1	3991	14 US-10-074-045-60	Sequence 60, Appl
44	57.6	8.1	6361	12 US-10-311-455-1114	Sequence 1114, Ap
45	57.6	8.1	6609	12 US-10-311-455-1855	Sequence 1855, Ap

ALIGNMENTS

RESULT 1
US-09-835-992A-19
; Sequence 19, Application US/09835992A
; Patent No. US20020037541A1
; GENERAL INFORMATION:
; APPLICANT: Obata, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AN
; FILE REFERENCE: L0461/7112
; CURRENT APPLICATION NUMBER: US/09/835,992A
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 08/896,164
; PRIOR FILING DATE: 1997-07-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (243)..(243)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (373)..(373)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (437)..(437)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (702)..(702)
; OTHER INFORMATION: n = a, c, g or t
US-09-835-992A-19

Query Match 99.4%; Score 710; DB 9; Length 714;
Best Local Similarity 100.0%; Pred. No. 7.9e-118;
Matches 714; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 CGCCAGAAAAGTATTATTATTCTATTAACATCTCTCAAGCATTTATTC 60

```
|||||
Db      1 CGCCAGAAAAAGTTATTTTAAATTTCTATTAAACATTCTTCTCAAGCATTATTTATCC 60
QY      61 TATATCTCACTGAATTTTAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATTA 120
Db      61 TATATCTCACTGAATTTTAAGAAATAACATTAGTATTAGAAAACTAGGAAAAAGATTA 120
QY      121 ATGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGAACGTTAT 180
Db      121 ATGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGAACGTTAT 180
QY      181 AAATGAAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATT 240
Db      181 AAATGAAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATT 240
QY      241 ATNCTTAATGATGGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT 300
Db      241 ATNCTTAATGATGGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT 300
QY      301 AAAAATATTTTATTTTAAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATT 360
Db      301 AAAAATATTTTATTTTAAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATT 360
QY      361 TTGATTTGGGAANAAAAATAGTCTTCTGATAGCATGAAATGCAAAATTTTATGATTTTA 420
Db      361 TTGATTTGGGAANAAAAATAGTCTTCTGATAGCATGAAATGCAAAATTTTATGATTTTA 420
QY      421 ATCTCACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAGTGACACACTA 480
Db      421 ATCTCACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAGTGACACACTA 480
QY      481 ATTACTGGCCAGCTGTGGCATGTGTTTCTTACTAGTCTCCCAAGGAAACTCTTA 540
Db      481 ATTACTGGCCAGCTGTGGCATGTGTTTCTTACTAGTCTCCCAAGGAAACTCTTA 540
QY      541 AATTGAATCTTCAGCAGATAATCCTTAATATATCTTGTAAAGCAAAACAAAGCTTTT 600
Db      541 AATTGAATCTTCAGCAGATAATCCTTAATATATCTTGTAAAGCAAAACAAAGCTTTT 600
QY      601 TGTTCACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTAC 660
Db      601 TGTTCACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTAC 660
QY      661 CCCAGACCATAATTACCATATTAACCTTGTGTTGACAGTGTGTTGCCAATTCA 714
Db      661 CCCAGACCATAATTACCATATTAACCTTGTGTTGACAGTGTGTTGCCAATTCA 714

RESULT 2
US-09-880-107-3296/c
; Sequence 3296, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3296
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U11313
```

```
US-09-880-107-3296
Query Match          94.2%; Score 672.8; DB 10; Length 1052;
Best Local Similarity 98.2%; Pred. No. 4e-111;
Matches 699; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY      2 GCCAGAAAAAGTTATTTTAAATTTCTATTAACAATCTCTCTCAAGCATTATTTATCCT 61
Db      1027 GCCAGAAAAAGTTATTTTAAATTTCTATTAACAATCTCTCTCAAGCATTATTTATCCT 968
QY      62 ATATCTCACTGAATTTTAAGAAATACATTAAGTATTAGAAAACTAGGAAAAAGATTA 121
Db      967 ATATCTCACTGAATTTTAAGAAATACATTAAGTATTAGAAAACTAGGAAAAAGATTA 908
QY      122 TGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGAACGTTATA 181
Db      907 TGCAGATAATTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGAACGTTATA 848
QY      182 AATGAAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATT 241
Db      847 AATGAAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATT 788
QY      242 TNCTTAATGATGGTGTTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTT 301
Db      787 TTCTTAATGATGGTGTTTTATGACTAATACACTGATTTTCAATGAAGAAACCCATGTT 728
QY      302 AAAATATTTTATTTTAAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATT 361
Db      727 AAAATATTTTATTTTAAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATT 668
QY      362 TGATTTGGGAANAAAAATACTGTTCTGATAGCATGAAATGCAAAATTTTATGATTTTAA 421
Db      667 TGATTTGGGAANAAAAATACTGTTCTGATAGCATGAAATGCAAAATTTTATGATTTTAA 608
QY      422 TCTCACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAGTGACACACTAA 481
Db      607 TCTCACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAGTGACACACTAA 548
QY      482 TTACTGGCCAGCTGTGGCATGTGTTTCTTACTAGTCTCCCAAGGAAAACTCTTAA 541
Db      547 TTACTGGCCAGCTGTGGCATGTGTTTCTTACTAGTCTCCCAAGGAAAACTCTTAA 489
QY      542 ATTGAATCTTCAGCAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTT 601
Db      488 ACTGAATCTTCAGCAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTT 429
QY      602 GTTTCACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACC 661
Db      428 GTTTCACATAGTCTTTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACC 370
QY      662 CCAGACCATAATTACCATATTAACCTTGTGTTGACAGTGTGTTGCCAATTTC 713
Db      369 CCAGACCATAATTACCATATTAACCTTGTGTTGACAGTGTGTTGCCAATTTC 318

RESULT 3
US-09-835-992A-20
; Sequence 20, Application US/09835992A
; Patent No. US20020037541A1
; GENERAL INFORMATION:
; APPLICANT: Obata, Yuichi
; APPLICANT: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING GASTRIC CANCER
; FILE REFERENCE: L0461/7112
; CURRENT APPLICATION NUMBER: US/09/835,992A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 08/896,164
; PRIOR FILING DATE: 1997-07-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 687
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (67)..(67)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (105)..(105)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (412)..(412)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (541)..(541)
; OTHER INFORMATION: n = a, c, g or t
; US-09-835-992A-20
```

Query Match 85.7%; Score 611.8; DB 9; Length 687;
Best Local Similarity 97.1%; Pred. No. 2.9e-100;
Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

```

QY 15 ATTTAATTTCTATTAACATCTCTCTCAAGCATTTTATCCATATCTCACTGAA 74
   |||||||
Db 1 ATTTAATTTCTATTAACATCTCTCTCAAGCATTTTATCCATATCTCACTGAA 60

QY 75 TTTTAAGAATAACATAGTATTAGAAAACTAGAAAAAAGATAATGACATAATTAA 134
   |||||||
Db 61 TTTTAANAAATAACATAGTATTAGAAAACTAGAAAAAAGATAATGACATAATTAA 120

QY 135 ACTTACATGAAAAAGAAAAATTATAACAAGGACTAGAACGTTATAATTGAAATGAGA 194
   |||||||
Db 121 ACTTACATGAAAAAGAAAAATTATAACAAGGACTAGAACGTTATAATTGAAATGAGA 180

QY 195 TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCATTATNCTTAATGATGG 254
   |||||||
Db 181 TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCATTATNCTTAATGATGG 240

QY 255 TGTATTGACTAATACTGATTTTTCAGAAGAAACCATGTTAAAAATATTTTAT 314
   |||||||
Db 241 TGTATTGACTAATACTGATTTTTCAGAAGAAACCATGTTAAAAATATTTTAT 300

QY 315 TTTAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTTATTTGATTTGGGAANA 374
   |||||||
Db 301 TTTAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTTATTTGATTTGGGAANA 360

QY 375 AAATACGTCTCTGATAGCATGAATGCAGAAATTTTATTTTATCTCACTAAATTTT 434
   |||||||
Db 361 AAATACGTCTCTGATAGCATGAATGCAGAAATTTTATTTTATCTCACTAAATTTT 420

QY 435 AANAACATTTGAGAAATGATTAATGACATGAAGTGCACAACATAATTACTGGCCAGCT 494
   |||||||
Db 421 AAGAACATTTGAGAAATGATTAATGACATGAAGTGCACAACATAATTACTGGCCAGCT 480

QY 495 GTTGACATTTGTTTCTTACTAGTTCTCTCCCAAGGAAACTCTTAATTTGAATCTTCAG 554
   |||||||
Db 481 GTTGACATTTGTTTCTTACTAGTTCTCTCCCAAGGAAACTCTTAATTTGAATCTTCAG 539

QY 555 CAGAATAATCCTTAATAATACTTTGTAAGCAAAAAGCTTTTGTGTTACATAGTTC 614
   |||||||
Db 540 CNGAATAA-CCTTAATAATACTTTGTAAGCAAAAAGCTTTTGTGTTACATAGTTC 597

QY 615 TTTGGGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACCACGACCATTAAT 674
   |||||||
Db 598 TTT-GGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACCACGACCATTAAT 656

QY 675 ACCATA-TTAACCTTTGTTNCGACAGTGT 704
   |||||||
Db 657 ACCCTATTTAACCTTTGTTNCGACAGTGT 687
```

RESULT 4
US-09-796-692-7228
; Sequence 7228, Application US/09796692
; Publication No. US20020198362A1

```

; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077,001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2000-03-17
; PRIOR FILING DATE: 2000-03-17
; PRIOR FILING DATE: 2000-04-27
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-04-28
; PRIOR FILING DATE: 2000-05-01
; PRIOR FILING DATE: 2000-05-01
; PRIOR FILING DATE: 2000-05-04
; PRIOR FILING DATE: 2000-05-04
; PRIOR FILING DATE: 2000-05-22
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-04
; PRIOR FILING DATE: 2000-08-04
; PRIOR FILING DATE: 2000-08-07
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7228
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-796-692-7228
```

Query Match 67.6%; Score 482.4; DB 10; Length 506;
Best Local Similarity 99.2%; Pred. No. 3.7e-77;
Matches 483; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 2 GCCAGAAAAAGTATTTTAAATTTCTATTAACATCTCTCAAGCATTTATCTCT 61
   |||||||
Db 20 GCCAGAAAAAGTATTTTAAATTTCTATTAACATCTCTCTCAAGCATTTATCTCT 79

QY 62 ATATCTCACTGAATTTTAAAGAAATACATTAAGTAACTAGGAAAAAGATAAA 121
   |||||||
Db 80 ATATCTCACTGAATTTTAAAGAAATACATTAAGTAACTAGGAAAAAGATAAA 139

QY 122 TGCAGATAATTTAACTTACATGAAAAAGAAAAATTATAACAAGGACTGAGAACGTTATA 181
   |||||||
Db 140 TGCAGATAATTTAACTTACATGAAAAAGAAAAATTATAACAAGGACTGAGAACGTTATA 199

QY 182 AATTGAATGAGATTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTA 241
   |||||||
Db 200 AATTGAATGAGATTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTA 259

QY 242 TNCCTAATGATGTTTATTAATGACTAATACTGATTTTTCAGAAGGAAACCATGTTA 301
   |||||||
Db 260 TNCCTAATGATGTTTATTAATGACTAATACTGATTTTTCAGAAGGAAACCATGTTA 319

QY 302 AAAATATTTTATTTTAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTT 361
   |||||||
Db 320 AAAATATTTTATTTTAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTT 379

QY 362 TGATTTGGGAAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATTTTAA 421
   |||||||
Db 380 TGATTTGGGAAANAATACTGTTCTGATAGCATGAATGCAAAATTTTATTTTAA 439
```


|||||
Db 301 AAATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTT 360
QY 363 GATTGGGAANAATACTGTTTCTGATGATGAATGCAAAATTTTAGATTTTAAAT 422
Db 361 GATTGGGAAGAAATACTGTTTCTGATGATGAATGCAAAATTTTAGATTTTAAAT 420
QY 423 CTCACATAATTTTAAANACTATGAGAAATGATTAATGACATGAAGTGACACAATAAT 482
Db 421 CTCACATAATTTTAAAGAACTATGAGAAATGATTAATGACATGAAGTGACACAATAAT 480
QY 483 TACTGG 488
Db 481 TACTGG 486
RESULT 7
US-09-835-992A-22
; Sequence 22, Application US/09835992A
; Patent No. US20020037541A1
; GENERAL INFORMATION:
; APPLICANT: Obata, Yutichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AN
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING GASTRIC CANCER
; FILE REFERENCE: L0461/7112
; CURRENT APPLICATION NUMBER: US/09/835,992A
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 08/896,164
; PRIOR FILING DATE: 1997-07-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (12)..(12)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (19)..(19)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (33)..(33)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (35)..(35)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (42)..(42)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (47)..(47)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (56)..(56)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (64)..(65)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (69)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (71)..(71)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (74)..(74)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (78)..(78)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure

LOCATION: (80)..(80)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (84)..(84)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (92)..(93)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (95)..(95)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (107)..(107)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (122)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (130)..(130)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (139)..(139)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (148)..(148)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (157)..(157)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (164)..(164)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (169)..(169)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (174)..(174)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (176)..(176)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (179)..(179)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (185)..(185)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (189)..(189)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (191)..(191)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (200)..(200)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (203)..(203)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (208)..(208)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (213)..(214)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (216)..(216)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (225)..(225)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
LOCATION: (229)..(229)

```
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (246)..(246)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (249)..(249)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (255)..(256)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (260)..(260)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (262)..(263)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (270)..(270)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (274)..(274)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (287)..(287)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (292)..(292)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (296)..(298)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (305)..(305)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (307)..(307)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (312)..(312)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (329)..(329)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (332)..(332)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (342)..(342)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (348)..(348)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (357)..(357)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (379)..(379)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (381)..(381)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (387)..(388)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (394)..(395)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (399)..(400)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (406)..(406)
```

```
Query Match      61.1%; Score 436.2; DB 9; Length 843;
Best Local Similarity 73.1%; Pred. No. 8e-69;
Matches 516; Conservative 0; Mismatches 187; Indels 3; Gaps 3;

QY      2 GCCAGAAAAAGTTATTTTAAATTTCTATTAACATTTCTCAAGCATTTATTTACCT 61
      2 GCCAAAAAANTTATTTAATTTCCCTATTANCNTCCCTCNCAANCAATTATTTNACCT 61
Db
QY      62 ATATCTCACTGAATTTTAAAGAAATAACATTAGTATTTAGAAAACTAGAAAAAAGATAAA 121
      62 ATNNCNCNCNGANTTTNANAANTACCTTTNNNTTAAAAAAAACCCTNGAAAAAATAAT 121
Db
QY      122 TGCAGATAATTAACTTTACATGAAAAAGAAAAATTATTAACAAAGACTGAGAACGTTATA 181
      122 NGCAATANTTTAACCTTNTCTTGAAAAANGAAATTTNTACCAANGACNGAANCNTTNTA 181
Db
QY      182 AATTGAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTATTTGTTCAATTA 241
      182 ATTNGAANTNAAATTATNTNGAANGCGCNCNCGAACAACCAANTTNATGTCCAATTA 241
Db
QY      242 TNCCTAATGATGTGTTTATAGACTAATAACACTGATTTTCAAGAGAAACCACTGTTA 301
      242 TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCNNNTTA 301
Db
QY      302 AAAATATTTTATTTTAAATAAAGCCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT 361
      302 AAANTNTTTTATTTTAAATAAACCNGTNTCCAAACCCNGATCANATTTCTTTNATTT 361
Db
QY      362 TGAATTTGGAAANAATACTGTTTCTGATAGCATGAAATGCAAAATTTTATGATTTTAA 421
      362 GGAATTTGGGAAAAAATNCNGTTCNNATACCNNGAANNCAANTTTTAAATTTTAA 421
Db
QY      422 TCTCACTAATTTTANAACATATTGAGAAATTGATTAATGACATGAAGTGCACACATAA 481
      422 CCCCCCTANTTTTAAANCTATNGAAAAANTNGATTANNGACCTGAATTGC-CAACCTAN 480
Db
QY      482 TTACTGGCCAGCTGTTGGCATTGTTCTTCTTACTTACTTCTCCCAAGGAAAACTCTTAA 541
      481 TTNCGGCCACCGTGGCGNTNGTNTTCTTACTTANTCCCCCAA-GGAANNCTTAA 539
Db
QY      542 ATTGAATCTTCAGCAGATAATCCTTAATATTAATCTTTGTAAGCAAAACAAGCTTTT 601
      540 NCGAANCTCCNCAAAATAACCTTAANTATCTTGGTAACCAACCAAAACCTTTT 599
Db
QY      602 GTTTACATAGTTCTTTGGGATTTTACTGTTCTTAATTTTATTTCTGAACCTCAATTTTACC 661
      600 GTTTACNTA-NTCCTTGGGATTTAACCGGTGCCCAATTTNATCCNGAACCAANTTTTCCC 658
Db
QY      662 CCAGACCATTAATTAACATTTAATTTGTTNTGACAGAGTTGTTGC 707
      659 CCNAACCATTAATTAACATTTTACCTTGGTAAGCNCAGTNGTTGC 704
Db

RESULT 8
US-10-102-524-161/c
; Sequence 161, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: THERAPY AND METHODS FOR THE
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 161
; LENGTH: 510
; TYPE: DNA
```

; ORGANISM: Homo sapiens
US-10-102-524-161

Query Match 26.3%; Score 188; DB 14; Length 510;
Best Local Similarity 96.4%; Pred. No. 1.7e-24;
Matches 213; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

OY 493 CTGTTGGCATTGTTTCTTACTTAGTCTCCCAAGGAAAACTCTTAATGATCTTC 552
|||||
DB 510 CTGTTGGCATTGTTTCTTACTTAGTCTCCCAA-GGAAAACTCTTAATGATCTTC 452
OY 553 AGCAGAATAATCCTTAATATATACTTTGTGAAGCAAAAGCTTTTGTGTACATAGT 612
|||||
DB 451 AGCAGAATAATCCTTAATATATACTTTGTGAAGCAAAAGCTTTTGTGTACATAGT 392
OY 613 TCTTTGGATTCTCTGTTCTAATTTTATCTGAACCTCAATTTACCAGCAGCATAA 672
|||||
DB 391 TCTTT-GGATTCTCTGTTCTAATTTTATCTGAACCTCAATTTACCAGCAGCATAA 333
OY 673 TTACCATATTAACTTTGTTNTGCACAGTTGTTGCCAATTC 713
|||||
DB 332 TTACCATATTAACTTTGTTATGACAGTTGTATGCAATTCC 292

RESULT 9

US-10-102-524-188/c

; Sequence 188, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:

; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-188

Query Match 26.3%; Score 188; DB 14; Length 510;
Best Local Similarity 96.4%; Pred. No. 1.7e-24;
Matches 213; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

OY 493 CTGTTGGCATTGTTTCTTACTTAGTCTCCCAAGGAAAACTCTTAATGATCTTC 552
|||||
DB 510 CTGTTGGCATTGTTTCTTACTTAGTCTCCCAA-GGAAAACTCTTAATGATCTTC 452
OY 553 AGCAGAATAATCCTTAATATATACTTTGTGAAGCAAAAGCTTTTGTGTACATAGT 612
|||||
DB 451 AGCAGAATAATCCTTAATATATACTTTGTGAAGCAAAAGCTTTTGTGTACATAGT 392
OY 613 TCTTTGGATTCTCTGTTCTAATTTTATCTGAACCTCAATTTACCAGCAGCATAA 672
|||||
DB 391 TCTTT-GGATTCTCTGTTCTAATTTTATCTGAACCTCAATTTACCAGCAGCATAA 333
OY 673 TTACCATATTAACTTTGTTNTGCACAGTTGTTGCCAATTC 713
|||||
DB 332 TTACCATATTAACTTTGTTATGACAGTTGTATGCAATTCC 292

RESULT 10

US-10-102-524-645
; Sequence 645, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:

; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 645
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-645

Query Match 26.3%; Score 188; DB 14; Length 510;
Best Local Similarity 96.4%; Pred. No. 1.7e-24;
Matches 213; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

OY 493 CTGTTGCATTGTTTCTTACTTAGTCTCCCAAGGAAAACTCTTAATGATCTTC 552
|||||
DB 1 CTGTTGCATTGTTTCTTACTTAGTCTCCCAA-GGAAAACTCTTAATGATCTTC 59
OY 553 AGCAGAATAATCCTTAATATATACTTTGTGAAGCAAAAGCTTTTGTGTACATAGT 612
|||||
DB 60 AGCAGAATAATCCTTAATATATACTTTGTGAAGCAAAAGCTTTTGTGTACATAGT 119
OY 613 TCTTTGGATTCTCTGTTCTAATTTTATCTGAACCTCAATTTACCAGCAGCATAA 672
|||||
DB 120 TCTTT-GGATTCTCTGTTCTAATTTTATCTGAACCTCAATTTACCAGCAGCATAA 178
OY 673 TTACCATATTAACTTTGTTNTGCACAGTTGTTGCCAATTC 713
|||||
DB 179 TTACCATATTAACTTTGTTATGACAGTTGTATGCAATTCC 219

RESULT 11

US-09-960-352-5962
; Sequence 5962, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5962
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB188-014-Q1-E1-G5
US-09-960-352-5962

Query Match 17.7%; Score 126.2; DB 10; Length 397;
Best Local Similarity 73.6%; Pred. No. 1.8e-13;
Matches 204; Conservative 0; Mismatches 60; Indels 13; Gaps 3;

OY 435 AANAATAATGAGAAATGATTAATGACATGAGTGCACACACTAATTAATGAGCACT 494
|||||
DB 14 AATTATCACTGAGAAATATATTAAGTCAATTAATGACACACACTAATTAATGAGCACT 73
OY 495 GTTGGCATTGTTTCTTACTTAGTCTCCCAAGGAAAACTCTTAATGATCTTCAG 554
|||||
DB 74 AATATGATAT-TCTGTTTCTTCATTCTCCCAAGGAAAC-----TTGAAATTTCGG 124

```
QY      555 CAGAAATATCCTTAATATATACTTTGTAAGCAAAACAAAGCTTTTGTTCATAGTTC 614
          |||||
Db      125 CAGAAATATCTTCAAAATATACATTAATAGCAAAATGAGAGCTTCTGTTT---ACATACT 180
QY      615 TTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACCAGACCATATTT 674
          |||||
Db      181 TTGTGATTTTGTCTATTTCTAACTTAATCTAAAACTCAATTTTACCAGACCATGATT 240
QY      675 ACCATATTAACCTTTGTTNGACAGTTGTTGCCAAT 711
          |||||
Db      241 ACCATATTAACCTTTGTAATGACAGTTGTTGCAACT 277
```

RESULT 12

```
US-09-960-352-13016/c
; Sequence 13016, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13016
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 56-BOVMS1-016-Q1-E1-F8
US-09-960-352-13016
```

```
Query Match      10.8%; Score 76.8; DB 10; Length 397;
Best Local Similarity 76.4%; Pred. No. 0.00012;
Matches 120; Conservative 0; Mismatches 33; Indels 4; Gaps 2;
```

```
QY      555 CAGAAATATCCTTAATATATACTTTGTAAGCAAAACAAAGCTTTTGTTCATAGTTC 614
          |||||
Db      397 CAGAAATAGCTTCAAAATATACATTAATAGCAAAATGAGAGC--TTCTGTTACATACTCA 340
QY      615 TTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACCAGACCATATTT 674
          |||||
Db      339 TT--GTATTCGTCTATTTATTAATCTTCTAAAACTCACTTTTACCCCAACCATGATT 282
QY      675 ACCATATTAACCTTTGTTNGACAGTTGTTGCCAAT 711
          |||||
Db      281 ACCATATTAACCTGTTGTAATGACAGATGTTGCAACT 245
```

RESULT 13

```
US-09-960-352-13260/c
; Sequence 13260, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13260
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 57-BOVMS1-016-Q1-E1-G2
US-09-960-352-13260
```

```
Query Match      10.5%; Score 75.2; DB 10; Length 449;
Best Local Similarity 85.6%; Pred. No. 0.00025;
Matches 83; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
QY      615 TTGGGATTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACCAGACCATATTT 674
          |||||
Db      415 TTGTGATTTTGTATTTCTTAACCTTTATTTCTAAAACTCAATTTTACCAGACCATGATT 356
QY      675 ACCATATTAACCTTTGTTNGACAGTTGTTGCCAAT 711
          |||||
Db      355 ACCATATTAACCTGTTAATGACAGTTGTTGCAACT 319
```

RESULT 14

```
US-10-198-846-6381/c
; Publication 6381, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6381
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 1, 8, 66, 71, 79, 81, 84, 85, 86, 88, 117, 119, 125, 134,
; LOCATION: 149, 156, 162, 166, 172, 179, 180, 194, 230, 249, 250, 253,
; LOCATION: 268, 271, 273, 274, 275, 276, 277, 286, 291, 299, 312, 333,
; LOCATION: 359, 366, 374, 393, 404, 406, 409, 411, 422, 424, 427
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 428, 433, 450, 465, 466, 468, 471, 487, 510, 512, 513, 516,
; LOCATION: 517, 518, 528, 530, 534, 539, 553, 555, 559, 575, 580, 588,
; LOCATION: 590, 610, 625, 626, 642, 645, 656, 658, 659, 675, 685, 686,
; LOCATION: 690, 700, 702, 706, 716, 731, 745, 748, 756, 798, 802
; OTHER INFORMATION: n = A,T,C or G
```

```
; NAME/KEY: misc_feature
; LOCATION: 807, 808, 819, 820, 821, 830, 838, 840, 844, 845, 849, 861,
; LOCATION: 871, 874, 875, 891, 892, 895, 898, 900, 901, 902, 904, 907,
; LOCATION: 917, 923, 925, 926, 927, 929, 955
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-6381
```

```
Query Match      9.5%; Score 67.6; DB 14; Length 960;
Best Local Similarity 41.2%; Pred. No. 0.0073;
Matches 217; Conservative 0; Mismatches 302; Indels 8; Gaps 1;
```

```
QY      67 TCACTGAATTTAAGAATAACATTAAGTAAAGAAAGTAGAAGAAAGATAATGACAG 126
          |||||
Db      930 TNANNANACCTANACCTTTAAANATNTNNANTTNAAANNAANNAATTTAAANTAN 871
QY      127 ATAATTAACCTTACATGAAAAAGGAAATTAACAAAGACTGAGAACGTTATAATG 186
          |||||
Db      870 AAATTTTNNAAAAAATAATTTAANNATANTNAATTTTNTAAAAAANNNAAAAATTA 811
QY      187 AAATGAGATTATTAATTTGAAAACCTGCATCTGAAGCAAACTTTATGTTCAATTATNCTT 246
          |||||
```

Db 810 AANNTTTNTAANTATTAACCAAAATTTTTTTTAAAAAAATTTTAAANNTAAT 751

QY 247 AATGATGGTGTTTTATGACTAATACACGTATTTTCAAGAAGAAACCAATGTTAAAAAT 306

Db 750 AANTNAATTTTTTAAATNAAAAAATTTNAAATTTAA-----NCAANTNT 699

QY 307 ATTTTATTTTAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATT 366

Db 698 TTTTNTTNTTNNATAAAAAANTTTTAAATAAAAAAANNANTTTTTTATNAAATA 639

QY 367 TGGGAANAAAAACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCA 426

Db 638 TTTAAAAAANNNTTTTNTTATANTTTAFAAAAAATTTTATTTNTTTTCTNA 579

QY 427 CTAATTTTAAANAATAATGAGAAATGATTAATGACATGAAGTGACAACTAATTACT 486

Db 578 AAANAAAAAATTTTNTNAAANTTTTAAAAACCTTNAATTNAAANAAATTTTTT 519

QY 487 GGCAGCTGTGGCATTGTGTTCTTACTTAAGTCTCCCAAGGAAAACTTAAATTGA 546

Db 518 NNNTNNANATAAAAAATTTATTTTNTTANNTCTATAAATTAANANANNAATATA 459

QY 547 ATCTTCAGACAGATAATCCTTAATATACTTTGTAAGCAAAACAAA 593

Db 458 ATAATTTTNNAAAAATAATTTTATANTTATNNANANAAATTTAAAA 412

RESULT 15

US-10-056-405-10

; Sequence 10, Application US/10056405

; Publication No. US20030166013A1

; GENERAL INFORMATION:

; APPLICANT: ZWIEBEL, LAURENCE J.

; TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF

; TITLE OF INVENTION: USE THEREOF

; FILE REFERENCE: N7841

; CURRENT APPLICATION NUMBER: US/10/056,405

; CURRENT FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: 60/264,649

; PRIOR FILING DATE: 2001-01-26

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 10

; LENGTH: 4985

; TYPE: DNA

; ORGANISM: Anopheles gambiae

US-10-056-405-10

Query Match 9.4%; Score 67.4; DB 12; Length 4985;

Best Local Similarity 44.4%; Pred. No. 0.014;

Matches 260; Conservative 0; Mismatches 325; Indels 0; Gaps 0;

QY 128 TAATTAACCTTACATGAAAAAGAAATTTATACAAGAGCTGAGAACGTTATAATTGA 187

Db 2277 TGATTTTATTTTAAAGAGAAAGATAAACCATTTTAGTAACCAATTAGTTACAGGAAC 2336

QY 188 AATGAGATTATAATTGAAAACTGCATCTGAAGCAAACTTATGTTCATATTATNCTTA 247

Db 2337 AAAATACAGAAATTATATTATTATTATTATTATTATTATTATTATTATTATTATT 2396

QY 248 ATGATGGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCCATGTTAAAAATA 307

Db 2397 TTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 2456

QY 308 TTTTATTTTTAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATT 367

Db 2457 TTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 2516

QY 368 GGGAANAAAACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCTCAC 427

Db 2517 TTATAATTATTACTTTATTATTATTATTATTATTATTATTATTATTATTATTATT 2576

QY 428 TAATTTTAAANAATAATGAGAAATTGATTAATGACATGAAGTGACAACTAATTACTG 487

Db 2577 TTATTATTATTATTATTATTATTATAATTAGATTATTATTATTATTATTATTATT 2636

QY 488 GCCAGCTGTGGCATTGTGTTCTTACTTAAGTCTCCCAAGGAAACCTTAAATTGAA 547

Db 2637 TTATAACAAATAATAATTATTATTATTATTATTATTATTAAATTAAATTTATTATTAA 2696

QY 548 TCTTCAGACAGATAATCCTTAATATACCTTGTAAAGCAAAACAAAGCTTTTGTGTTAC 607

Db 2697 TTATTATTATTGTATTTCATTATTATACATTATTATCATATAATTAATTTTATTATGATT 2756

QY 608 ATAGTCTTGGGATTTTACTGTTCCTAATTTTATTCTGAAGCTCAATTTTACCCAGAC 667

Db 2757 ATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 2816

QY 668 CATTAATTACCATATTACTTTGTTNCGACAGTTGTTGCCAATT 712

Db 2817 ATTATTATTATTATTAAATATTATTATTATTATTATTATTATTATTATTATTATT 2861

Search completed: November 27, 2003, 12:52:53

Job time : 275.238 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 27, 2003, 06:00:22 ; Search time 52.701 Seconds
(without alignments)
5979.910 Million cell updates/sec

Title: US-09-835-992A-19
Perfect score: 714
Sequence: 1 CGCCAGAAAAGTATTCTTCAATTAACATTTAGTATTGAAAACTAGAAAAAGATTA

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	710	99.4	714	3	US-08-896-164-19 Sequence 19, Appl
2	611.8	85.7	687	3	US-08-896-164-20 Sequence 20, Appl
3	436.2	61.1	843	3	US-08-896-164-22 Sequence 22, Appl
4	165.6	23.2	173	4	US-09-313-294A-1553 Sequence 1553, Ap
5	66.4	9.3	5852	1	US-07-867-106-2 Sequence 2, Appli
6	61.2	8.6	20674	4	US-09-641-638-651 Sequence 651, App
7	57.4	8.0	927	4	US-09-134-001C-150 Sequence 150, App
8	55.2	7.7	837	3	US-08-998-416-288 Sequence 288, App
9	53	7.4	19124	2	US-08-487-826B-13 Sequence 13, Appl
10	52.4	7.3	615	3	US-08-998-416-186 Sequence 186, App
11	51	7.1	19124	2	US-08-487-826B-13 Sequence 13, Appl
12	50.2	7.0	678	1	US-07-991-867B-23 Sequence 23, Appl
13	50.2	7.0	678	1	US-08-107-755A-23 Sequence 23, Appl
14	50.2	7.0	678	2	US-08-544-332-23 Sequence 23, Appl
15	50.2	7.0	678	4	US-09-370-861A-23 Sequence 23, Appl
16	50.2	7.0	1788	4	US-09-601-198-101 Sequence 101, App
17	50.2	7.0	6124	4	US-08-213-419B-3 Sequence 3, Appli
18	50.2	7.0	6768	1	US-08-107-755A-1 Sequence 1, Appli
19	50.2	7.0	8457	1	US-07-991-867B-1 Sequence 1, Appli
20	50.2	7.0	8457	2	US-08-544-332-1 Sequence 1, Appli
21	50.2	7.0	8457	4	US-09-370-861A-1 Sequence 1, Appli
22	50	7.0	3232	4	US-09-333-214-3 Sequence 3, Appli
23	49.8	7.0	6243	2	US-09-056-075-1 Sequence 1, Appli
24	49.8	7.0	20674	4	US-09-641-638-651 Sequence 651, App
25	49.2	6.9	1692	4	US-09-601-198-63 Sequence 63, Appl
26	48.8	6.8	3234	1	US-08-286-325A-7 Sequence 7, Appli
27	48.6	6.8	615	3	US-08-998-416-186 Sequence 186, App

28	47.8	6.7	6124	4	US-08-213-419B-3 Sequence 3, Appli
29	47.6	6.7	711	4	US-09-601-198-111 Sequence 111, App
30	47.4	6.6	5340	4	US-09-627-122-21 Sequence 21, Appl
31	47.4	6.6	6152	3	US-08-973-462-1 Sequence 1, Appli
32	46.8	6.6	1511	1	US-07-991-867B-8 Sequence 8, Appli
33	46.8	6.6	1511	1	US-08-107-755A-8 Sequence 8, Appli
34	46.8	6.6	1511	2	US-08-544-332-8 Sequence 8, Appli
35	46.8	6.6	1511	4	US-09-370-861A-8 Sequence 8, Appli
36	46.8	6.6	2435	3	US-09-306-593-1 Sequence 1, Appli
37	46.2	6.5	827	3	US-08-998-416-535 Sequence 535, App
38	46	6.4	240	1	US-08-628-417-6 Sequence 6, Appli
39	46	6.4	658	3	US-08-998-416-595 Sequence 595, App
40	46	6.4	5852	1	US-07-867-106-2 Sequence 2, Appli
41	45.8	6.4	10640	4	US-09-417-485D-5 Sequence 5, Appli
42	45.6	6.4	636	3	US-08-998-416-1137 Sequence 1137, Ap
43	45.6	6.4	1474	3	US-08-821-994-64 Sequence 64, Appl
44	45.6	6.4	2435	3	US-09-306-593-1 Sequence 1, Appli
45	45.2	6.3	945	4	US-09-601-198-177 Sequence 177, App

ALIGNMENTS

RESULT 1
US-08-896-164-19
; Sequence 19, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-896-164-19
Query Match 99.4%; Score 710; DB 3; Length 714;
Best Local Similarity 100.0%; Pred. No. 3.2e-132;
Matches 714; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCAGAAAAGTATTCTTCTAATTAACATTTCTCAAGCATTTATTC 60
Db 1 CGCCAGAAAAGTATTCTTCTAATTTCTAATTAACATTTCTCAAGCATTTATTC 60
QY 61 TATATCTCACTGAATTTTAAGAATAACATTTAGTATTGAAAACTAGAAAAAGATTA 120

```
Db      61 TATATCTCACTGATTTTAAAGAAATAACATTAGTATTGAAAAAAGTAA 120
QY      121 ATGCAGATAATTAACTTACATGAAAAAGAAATTATTAACAAGAGCTGAGACGTTAT 180
Db      121 ATGCAGATAATTAACTTACATGAAAAAGAAATTATTAACAAGAGCTGAGACGTTAT 180
QY      181 AAATTGAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATT 240
Db      181 AAATTGAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATT 240
QY      241 ATNCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAGAAACCCATGTT 300
Db      241 ATNCTTAATGATGGTGTGTTTATGACTAATACACTGATTTTCAAGAGAAACCCATGTT 300
QY      301 AAAAATATTTTATTTTAAATAAGCCCTGTGTCAAGCTCTGATCATATTCTTTATT 360
Db      301 AAAAATATTTTATTTTAAATAAGCCCTGTGTCAAGCTCTGATCATATTCTTTATT 360
QY      361 TTGATTTGGGAANAAAAATACCTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTTA 420
Db      361 TTGATTTGGGAANAAAAATACCTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTTA 420
QY      421 ATCTCACTAATTTTAAANAACATGAGAAATGATTAAATGACATGAAGTGACACACACTA 480
Db      421 ATCTCACTAATTTTAAANAACATGAGAAATGATTAAATGACATGAAGTGACACACACTA 480
QY      481 ATTACTGGCCAGCTGTGTGGCATGTGTTCTTACTAGTCTCCCAAGGAAAACTCTTA 540
Db      481 ATTACTGGCCAGCTGTGTGGCATGTGTTCTTACTAGTCTCCCAAGGAAAACTCTTA 540
QY      541 AATTGAATCTTCAGAGAAATAATCCTTAAATATATCTTTGTAAGCAAAACAAAGCTTTT 600
Db      541 AATTGAATCTTCAGAGAAATAATCCTTAAATATATCTTTGTAAGCAAAACAAAGCTTTT 600
QY      601 TGTTTACATAGTCTTTGGGATTTTACTCTGTTCTAATTTTATTCTGAAACTCAATTTC 660
Db      601 TGTTTACATAGTCTTTGGGATTTTACTCTGTTCTAATTTTATTCTGAAACTCAATTTC 660
QY      661 CCCAGACCATAATTACCATATTAACCTTGTNTGCACAGTTGTTGCCAATTCA 714
Db      661 CCCAGACCATAATTACCATATTAACCTTGTNTGCACAGTTGTTGCCAATTCA 714

RESULT 2
US-08-896-164-20
; Sequence 20, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
```

```
REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-896-164-20

Query Match      85.7%; Score 611.8; DB 3; Length 687;
Best Local Similarity 97.1%; Pred. No. 8e-113;
Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

QY      15 ATTTAATTTTCTATTAAACATTTCTTCAAGCATTTATTTATCTATATCTCACTGAA 74
Db      1 ATTTAATTTTCTATTAAACATTTCTTCAAGCATTTATTTATCTATATCTCACTGAA 60
QY      75 TTTAAGAAATAACATTAGTATTAGAAAAAAGTAAAGAAAAAGATAAATGAGATAATTAA 134
Db      61 TTTAANAATAACATTAGTATTAGAAAAAAGTAAAGAAAAAGATAAATGAGATAATTAA 120
QY      135 ACTTACATGAAAAAGAAAAATTTATACAAAGAGCTGAGAACGTTATTAATGAAATGAGA 194
Db      121 ACTTACATGAAAAAGAAAAATTTATACAAAGAGCTGAGAACGTTATTAATGAAATGAGA 180
QY      195 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG 254
Db      181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATNCTTAATGATGG 240
QY      255 TGTTTATGACTAATACACTGATTTTTCAGAGAGAAACCCATGTTAAAAATATTTTAT 314
Db      241 TGTTTATGACTAATACACTGATTTTTCAGAGAGAAACCCATGTTAAAAATATTTTAT 300
QY      315 TTTAAAAATAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTTTGAATTGGGAANA 374
Db      301 TTTAAAAATAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTTTGAATTGGGAANA 360
QY      375 AAATACGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAATCTCAATAATTT 434
Db      361 AAATACGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAATCTCAATAATTT 420
QY      435 AANAACATTTGAGAAATGATTATATGACATGAGCTGCACACACTAATTTACTGGCCAGCT 494
Db      421 AAGAACATTTGAGAAATGATTATATGACATGAGCTGCACACACTAATTTACTGGCCAGCT 480
QY      495 GTTGCAATGTGTTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTAATTTGAATCTTCAG 554
Db      481 GTTGCAATGTGTTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTAATTTGAATCTTCAG 539
QY      555 CAGATAATCCTTAATATATACCTTTGTAAGCAAAACAAAGCTTTTGTGTTACATAGTTC 614
Db      540 CNGAATAA-CCTTAATATATACCTTTGTAAGCAAAACAAAGCTTTTGTGTTACATAGTTC 597
QY      615 TTTGGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACCCGAGACATAATT 674
Db      598 TTT-GGATTTTACTGTTCTTAATTTTATTTCTGAAACTCAATTTTACCCGAGACATAATT 656
QY      675 ACCATA-TTAACTTTGTTNTGCACAGTTGTT 704
Db      657 ACCCTATTAACTTTGTTATGACACAGTTGTT 687

RESULT 3
US-08-896-164-22
; Sequence 22, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
```

```

1 TITLE OF INVENTION:  DIAGNOSING AND TREATING GASTRIC CANCER
2
3 NUMBER OF SEQUENCES:  87
4
5 CORRESPONDENCE ADDRESS:
6
7 ADDRESSEE:  Felde & Lynch
8
9 STREET:  805 Third Avenue
10
11 CITY:  New York City
12
13 STATE:  New York
14
15 COUNTRY:  USA
16
17 ZIP:  10022
18
19 COMPUTER READABLE FORM:
20
21 MEDIUM TYPE:  Diskette, 3.50 inch, 1.44mb
22
23 COMPUTER:  IBM PS/2
24
25 OPERATING SYSTEM:  PC-DOS
26
27 SOFTWARE:  wordperfect
28
29 CURRENT APPLICATION DATA:
30
31 APPLICATION NUMBER:  US/08/896,164
32
33 FILING DATE:  July 17, 1997
34
35 CLASSIFICATION:  424
36
37 ATTORNEY/AGENT INFORMATION:
38
39 NAME:  No. 6218521man D. Hanson
40
41 REGISTRATION NUMBER:  30,946
42
43 REFERENCE/DOCKET NUMBER:  LUD 5499 - JEL/NDH/SLH
44
45 TELECOMMUNICATION INFORMATION:
46
47 TELEPHONE:  (212) 688-9200
48
49 TELEFAX:  (212) 838-3884
50
51 INFORMATION FOR SEQ ID NO:  22:
52
53 SEQUENCE CHARACTERISTICS:
54
55 LENGTH:  843 base pairs
56
57 TYPE:  nucleic acid
58
59 STRANDEDNESS:  double
60
61 TOPOLOGY:  linear
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971

```

Query Match	61.1%;	Score 436.2;	DB 3;	Length 843;
Best Local Similarity	73.1%;	Pred. No. 4e-78;		
Matches 516;	Conservative	0;	Mismatches 187;	Indels 3;
				Gaps 3;

OY	2	GCAGAAAAAGTATTTTAAATTTCTATTAACAATCTCTCTCAAGCATTATTTATCCT	61
Db	2	GCCAAAAAANTTATTTNAATTTCTATTAANCNTCTCCNCAAAACATTTATTAACCT	61
OY	62	ATATCTCACTGAATTTTAAGAATAACATAGTATTTAGAAAACTAGAAAAAAGATAAA	121
Db	62	ATNNCNCNCNGANTTTNAAAAANTACCTTNNNTNTTAAAAAACCTNGAAAAAAATAAT	121
OY	122	TGCAGATAATTTAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGACGTATA	181
Db	122	NGCAAATANTTTAACCTTCTGAAAAANGAAATTTNTATCCAANGACNGAAANCNTNTA	181
OY	182	AATTGAAATGAGATTATTAATTTGAAAACTGCACTGAAAAGCAAACTTATTTGTTCAATTA	241
Db	182	ATTNGAANTNAATTAATTANTTNGAANGCGCNCNCGAAGAACCAANTTNATGTCCTCAATTA	241
OY	242	TNCTTAATGATGCTGTTTATGACTAATACTGATTTTTCAGAAGAAACCACTGTTA	301
Db	242	TCCTNAANGAGGNNTTTNANNACTAATNCCNGAATTTTCCAATANGGAANCCNNNTTA	301
OY	302	AAAAATATTTTATTTTAAAAAATAAGCCTGTGTTCAGAGCTCTGATCATATTTCTTTATTT	361
Db	302	AAANTNTTTTNAATTTTAAAAATPACCNGTNTCCAACCCNGATCANATTCCTTNAATTT	361
OY	362	TGATTTGGGAANAAAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTATAGATTTTAA	421
Db	362	GGATGGGGGAAAAAAATNCNGTTCNNATPACCNNGAANNGCAAANTTTTAAATTTTAA	421
OY	422	TCTCACTAATTTTAANAACATATGAGAAATGATTAATGACATGAAGTGCACAACACTAA	481
Db	422	CCCCCTTANTTTTAAAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCTAN	480
OY	482	TTACTGGCCACGCTGTGGCATGTGTCTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTA	541
Db	481	TTNCNGGCCACCGTGGCGTNGTNTTCTTACTTANTTCCCCCCTAA-GGAAANNCCTTAA	539

[illegible]

```

RESULT 4
US-09-313-294A-1553/C
: Sequence 1553, Application US/09313294A
: Patent No. 6476212
: GENERAL INFORMATION:
: APPLICANT: Lalgudi, Raghunath V.
: APPLICANT: Ito, Laura Y.
: APPLICANT: Sherman, Bradley K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
: FILE REFERENCE: PL-0017 US
: CURRENT APPLICATION NUMBER: US/09/313,294A
: CURRENT FILING DATE: 1999-05-14
: NUMBER OF SEQ ID NOS: 7600
: SOFTWARE: PERL Program
: SEQ ID NO 1553
: LENGTH: 173
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. 6476212 700551132H1
US-09-313-294A-1553

```

Query Match	23.2%	Score 165.6	DB 4	Length 173
Best Local Similarity	97.7%	Pred. No. 9.7e-25		
Matches 168	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	8	AAAGTTATTTTAAATTTCTAATTAAACATTCCTCTCAAGCATTTATTTTATCCTATATCT	67	
Db	172	AAAAAGTTATTTAATTTCTCAATTAAACATTCCTCTCAAGCATTTATTTTATCCTATATCT	113	
QY	68	CACTGAATTTTAAAGAAATACATTAGTATTAGAAAACTAGAAAAAAGATAAATGCAGA	127	
Db	112	CACTGAATTTTAAAGAAATACATTAGTATTAGAAAACTAGAAAAAAGATAAATGCAGA	53	
QY	128	TAAITTAACCTTACATGAAAAAGAAAAATTATACCAAGGACTGAGAACGTTA	179	
Db	52	TAAITTAACCTTACATGAAAAAGAAAAATTATACCAAGGACTGAGAACGTTA	1	

```

RESULT 5
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostellum
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; NAME/KEY: CDS
; LOCATION: 2378..5038
; US-07-867-106-2

Query Match          9.3%; Score 66.4; DB 1; Length 5852;
Best Local Similarity 46.8%; Pred. No. 5.3e-05;
Matches 240; Conservative 0; Mismatches 269; Indels 4; Gaps 1;

QY 16 TTTTAATTTCTATTAACAATCTCTCAAGCATTAATTTTATCCTATATCTCACTGAAT 75
Db 5751 TTGTATTTTATATATAGTTAATTTGTTGTTTCTTAACTTATATTTCTAATTTTGA 5692

QY 76 TTTAAGAATTAACATTAAGTATTAAGAAACTAGAGAAAGATAATGCAGATAATTAA 135
Db 5691 TTTATAAAATTAATAATTAATTTTAAATAAATAAATAAATAAATAAATAAATAA 5632

QY 136 CTTACATGAAAAAGAAAAATTATACAAGAGCTGAGAACGTTATAATTGAATGAGAT 195
Db 5631 ATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5572

QY 196 TATTAATTTGAAACTGCACTGGAAGCAACTTATGTTCAATTATNCTTAATGATGT 255
Db 5571 TAAAAAATGTAAGGGTTTTTTTAAATATGATGATTTTATTTTAAATCATTTGA 5512

QY 256 GTTTTATGACTAATACACTGATTTTTCAGAAGAGAAACCATGTAAAAATATTTTAT 315
Db 5511 CGAGATTAAAAAATCTTAAAACAATAAACATATTTGATTTTCTTTTCTTTTCTTT 5452

QY 316 TT---AAAAATAGCCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTTGGA 371
Db 5451 TTTTAAAAATTCAAATATAATAATTAATTAATCTATTAATCTGATGAACTTCAATTT 5392

QY 372 ANAAAAATAGTCTTCTGATAGCATGAATGCAAAATTTTGAATTTTAAATCTCACTAAT 431
Db 5391 TAAATTTAATATATTTAATTTTAAAAAATTAGTATCTATCATATATATTTCCATG 5332

QY 432 TTTAANAATATGAGAAATGATTAATGACATGAGTGCACACACTAATTAATGAGCA 491
Db 5331 TTTTAAAAATTTTAAAGTATTTTAAATTAATGTAAGTACGACGAGATTAATATGGA 5272

QY 492 GCTGTTGGCATGTGTTCTTACTAGTCTCC 524
Db 5271 AATGATAGTACGATCCTTTTCCCAAGTTTCC 5239
```

```

RESULT 6
US-09-641-638-651/c
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
```

```

LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc feature
LOCATION: 17555..20674
OTHER INFORMATION: 3' regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1570
OTHER INFORMATION: 10-509-295 : deletion of C
NAME/KEY: allele
LOCATION: 1827
OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTTTT
NAME/KEY: allele
LOCATION: 2048
OTHER INFORMATION: 10-511-62 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2323
OTHER INFORMATION: 10-511-337 : insertion of T
NAME/KEY: allele
LOCATION: 2341
OTHER INFORMATION: 10-512-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 2623
OTHER INFORMATION: 10-512-318 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2832
OTHER INFORMATION: 10-513-250 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2844
OTHER INFORMATION: 10-513-262 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2934
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338

```

OTHER INFORMATION: 10-347-74	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 6375		
OTHER INFORMATION: 10-347-111	:	polymorphic base G or C
NAME/KEY: allele		
LOCATION: 6429		
OTHER INFORMATION: 10-347-165	:	polymorphic base C or T
NAME/KEY: allele		
LOCATION: 6467		
OTHER INFORMATION: 10-347-203	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 6484		
OTHER INFORMATION: 10-347-220	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 6534		
OTHER INFORMATION: 10-347-271	:	polymorphic base A or T
NAME/KEY: allele		
LOCATION: 6611		
OTHER INFORMATION: 10-347-348	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 7668		
OTHER INFORMATION: 10-348-391	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 8608		
OTHER INFORMATION: 10-349-47	:	polymorphic base C or T
NAME/KEY: allele		
LOCATION: 8658		
OTHER INFORMATION: 10-349-97	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 8703		
OTHER INFORMATION: 10-349-142	:	polymorphic base G or C
NAME/KEY: allele		
LOCATION: 8777		
OTHER INFORMATION: 10-349-216	:	deletion of CTG
NAME/KEY: allele		
LOCATION: 8785		
OTHER INFORMATION: 10-349-224	:	polymorphic base G or T
NAME/KEY: allele		
LOCATION: 8926		
OTHER INFORMATION: 10-349-368	:	polymorphic base C or T
NAME/KEY: allele		
LOCATION: 12171		
OTHER INFORMATION: 10-350-72	:	polymorphic base C or T
NAME/KEY: allele		
LOCATION: 12429		
OTHER INFORMATION: 10-350-332	:	polymorphic base C or T
NAME/KEY: allele		
LOCATION: 13341		
OTHER INFORMATION: 10-507-170	:	polymorphic base A or G
NAME/KEY: allele		
LOCATION: 13492		
OTHER INFORMATION: 10-507-321	:	polymorphic base A or C
NAME/KEY: allele		
LOCATION: 13524		
OTHER INFORMATION: 10-507-353	:	polymorphic base C or T
NAME/KEY: allele		
LOCATION: 13535		

Query Match	8.6%	Score 61.2;	DB 4;	Length 20674;
Best Local Similarity	47.6%	Pred. No. 0.00064;		
Matches 207; Conservative	0;	Mismatches 226;	Indels 2;	Gaps 1;

QY	9	AAAGTATTTTAATTCTATTAAACATCTCTCAAGCATTATTTATCCATATATCTC	68
Db	11513	AAATTTAAATTAATTATATAATTAATTTTAAATTTAAATTTAATTTAATATT	11454
QY	69	ACTGAATTTTAGAATAACATTAGATTGAAAAACTAGAAAAAAGATTAATGCAGAT	128
Db	11453	A--AAATAAAAATATTAAAAATTTAATTAATATTGAAGCAATTAAAAATTAAATT	11396
QY	129	AATTAACCTTACATGAAAAAGGAAAATTATAACAAGAGCTGACAACGTTAATTTGAA	188
Db	11395	AATTAATATTAAATTTAAATTTAAATTTAATATTAAAAATTTAAATTTAAATTTAA	11336

QY	189	ATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACCTTATTTGTTCAATTATNCTTAA	248
Db	11335	TTTAATATTTAAATTAATTAAATTAAGTTAAATTAATATTTAAATTTAAATTAAGTTAAATTA	11276
QY	249	TGATGCTGTTTATGACTAATACACTGATTTTTCAGAAGGAAACCATGTTAAAAATAT	308
Db	11275	ATATTAAATTAATTAATTAAATTAATATTAATAAATTAATAATGTTTATTTAAAAATTAA	11216
QY	309	TTTATTTTAAAAAATAAGCCTGTTCAAGCTCTGATCATATTTCTTTATTTGATTTG	368
Db	11215	ATTTAAATTTTAAATTTTAAATTTATTTAAATTTTAAATTTTAAATTTTAAATTTT	11156
QY	369	GGAAAAAATACTGTTTCTGATAGCATGAATGCAAATTTTATGATTTTAACTCTCACT	428
Db	11155	AATTTAATATTAATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTA	11096
QY	429	AATTTTAAANAATAAT 443	
Db	11095	ATATTAAAAATATTTT 11081	

```

RESULT 7
US-09-134-001C-150/c
; Sequence 150, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 150
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-150

```

Query Match	8.0%;	Score 57.4;	DB 4;	Length 927;
Best Local Similarity	49.7%;	Pred. No. 0.0027;		
Matches 171; Conservative	0;	Mismatches 169;	Indels 4;	Gaps 1;

OY	214	TCTGAAGCAAACTTTATTGTTCAATTA	INCTTAATGATGGTGTTCATGACTATAAC	273
Dd	606	TATAATGACTTATTTTCGTTAATAITC	GTTTAAAGTGCTCTTTGAGATAAATCAC	547
OY	274	TGATTTTTCAAGAAGAAACCAGTGTAAA	AATAATTTTATTTTAAAAATAAGCCGTGT	333
Dd	546	TAAFTTTTCATTAAAGATATTTGCGAA	ATAAGATCTTATCGTTAATATCCAGATAT	487
OY	334	TCAAGCTCTGATCATATTTCTTTATT	TGATTGGGAANNAATACT---GTTTCTGA	389
Dd	486	AAAAGCACAATCCATTTCTCTATCTT	CATGTTCAATTAATTATACATAATCTGT	427
OY	390	TAGCATGAATGCAAAAATTTTAGATTT	TTAATCTCACTAATTTAANAACATAATGAGA	449
Dd	426	TTTTAAATTAATTTGAATTTTGAAAT	TTTCCCTTATATTCATTCAATAATTTAGSTA	367
OY	450	ATTGATTAAATGACATGAAGTGCACA	CACTAATTACTGGCCAGCTGTGGCATGTGTTT	509
Dd	366	ATGAACGACGCTATTTGTTCAATTG	TCCTATGTTAATATTTAATTTTATTGTTGCT	307
OY	510	CTTACTTAGTTCTCCCAAAGGAAAACT	CTTAAATTTGAATCTTCA 553	
Dd	306	TACATTATGTTCTGCTTCATGTAATA	ATTGTAATATTTCTCTTAA 263	

RESULT 8
US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:

APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Rebischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentm Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1241RP
US-08-998-416-288

```

Query Match	7.7%;	Score 55.2;	DB 3;	Length 837;
Best Local Similarity	46.4%;	Pred. No. 0.0072;		
Matches 212; Conservative	0;	Mismatches 241;	Indels 4;	Gaps 1;

[illegible]

Db 339 ATTATTTTAAACAATTAAATAA---TAAATATAATTAATTATGTGATAACTATTTAAT 284

QY 305 ATATTTTATTTTAAAAAATAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGA 364

Db 283 AATTATTAAAGAAAATAAATATATCTAATAATATTTTAACTAATTTAAATTTGAA 224

QY 365 TTGGGAANAAATACTGTTCTGTGATGACGAAATGCAAAATTTTGTAGATTTTAACT 424

Db 223 CATAGACTAAATAGTATTCATATTAATATATTTTAAATATATATAATTAATAAT 164

QY 425 CACTAATTTTAAANAATAATGAGAAATGATTAATGA 461

Db 163 GATGATTAAGTAAATTAATAATAATTATATATAATGA 127

RESULT 9

US-08-487-826B-13

Sequence 13, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 19124 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-487-826B-13

Query Match 7.4%; Score 53; DB 2; Length 19124;

Best Local Similarity 45.7%; Pred. No. 0.026;

Matches 215; Conservative 0; Mismatches 253; Indels 2; Gaps 1;

QY 6 GAAAAAGTTATTTTAAATTTCTATTAACATTTCTCAAGCATTTATTTATCTATAT 65

Db 15367 GATGTAAGTGTATGTTTATATATATATTTTATATATATGATTTATTAATAAAGA 15426

QY 66 CTCACCTGAATTTTAAGAAATAACATTAAGTATTAGAAAAAAGTAAATGCA 125

Db 15427 AATATAAAACAATTTTATTAATAATGAAAAAAGAAAAATGAATATATAAAAAATTTA 15486

QY 126 GATAATTAACTTACATGAAAAAGAAATTAACAAGAGACTGAGAACGTTATAATT 185

Db 15487 TTAATAATAAAAAAGAAAAAAGAAAAAAGAAAAAATTTTAAAAAAATAAATAAAT 15546

QY 186 GAAATGAGATTAAATTTTGAAGAACTGCATCGAAGCAAACTTATTTGTTCAATTATNCT 245

Db 15547 ATAATAAATAATTAATTTTGATA--GAATAAATAATGAAAAAGATTATCAAAAAAAT 15604

QY 246 TAATGATGTTTAACTAATAACACTGATTTTCAAGAGAAAGCAACGTTAAAAA 305

Db 15605 AAAAAAATTTTATATAAAAAAATGATTAATAAAAAAATAAAAACAAGAAAGAAA 15664

QY 306 TATTTTATTTTAAATAAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGAT 365

Db 15665 AAAAAAACATTAATAAAAAAATATATATATCAATAAAAAACAAGAAAAAAT 15724

QY 366 TTGGGAANAAATACTGTTCTGTGATGACATGAATGCAAAATTTTAGATTTTAACTC 425

Db 15725 ATATTAATAATAATAATATATATCATTAATAATAAAAAAATTAATAAATAA 15784

QY 426 ACTAATTTTAAANAATAATGAGAAATGATTAATGACATGAAGTGACAA 475

Db 15785 AATATATACATAAATAAATAAATAATTTATTTAAATAAATAAATAAATAA 15834

RESULT 10

US-08-998-416-186

Sequence 186, Application US/08998416

Patent No. 6239264

GENERAL INFORMATION:

APPLICANT: Philippsen, Peter

APPLICANT: Pohlmann, Rainer

APPLICANT: Steiner, Sabine

APPLICANT: Mohr, Christine

APPLICANT: Wendland, Jurgen

APPLICANT: Knechtle, Philipp

APPLICANT: Reibischung, Corinne

TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPI

TITLE OF INVENTION: AND USES THEREOF

NUMBER OF SEQUENCES: 1152

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6239264artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: No. 6239264th Carolina

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/998,416

FILING DATE: 24-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: CH 0016/97

FILING DATE: 31-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 186:

SEQUENCE CHARACTERISTICS:

LENGTH: 615 base pairs

TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1074RP
US-08-998-416-186

Query Match 7.3%; Score 52.4; DB 3; Length 615;
Best Local Similarity 47.0%; Pred. No. 0.025;
Matches 225; Conservative 0; Mismatches 249; Indels 5; Gaps 2;

QY 5 AGAAAAAGTATTATTCTAATTAACATCTCTCAAGCATTATTATCTATA 64
Db |||||
26 ATAAAGATTAACTTTTATTATATATTAAATTAATTTAACTATT 85
QY 65 TCTCACTGAATTTAAGAAATACATTAGTATTAGAAAACTAGAAAAAGATAATGC 124
Db |||||
86 ATTATCATTTTAAATAATTAATTATTGATTATTAATCT---TATTATATAATTAT 141
QY 125 AGATAATTAACTTACATGAAAAAGAAATTTATAACAAGACTGAGAACGTTAATAAT 184
Db |||||
142 TATATAATTACTTAATTCATCATTTAATATTATTATTAATAAATAATATTATA 201
QY 185 TGAATAGAGATTATATTGAAAACTGCATCTGAAAGCAACTTTATTGTTCAATTATNC 244
Db |||||
202 TATGAATACATTATTAGTCT-ATGTTCAAATTTAAATTAGTTATTAAAAATATTAGAT 260
QY 245 TTAATGATGTTGTTTATGACTAATACACTGATTTTCAAGAGAAACCAGTTAAAA 304
Db |||||
261 ATTATTTATTTCTTAATAAATTATTAAATAGATTATCAATAATTATATTATTATT 320
QY 305 ATATTTTATTTTAAAAAAGCCTGTGTTCAAGCTTGATCATATTTCTTTATTTTGA 364
Db |||||
321 AATTGTTTATTAATAATTAATTTTATTATTATAAAGATTAAATTATTAAATTTGT 380
QY 365 TTGGAANAATACTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACT 424
Db |||||
381 AAATTTATTTTATTTATTAATAATCTATTTTATAAATAATTATGTTGATTATTAATT 440
QY 425 CACTAATTTTAAANACTATTGAGAAATTGATTATGACATGAAGTGACACACTAATT 483
Db |||||
441 AACTTTTATAGAATTATTATAAATAATTAACTTTAATTCTTATTATTAAATT 499

RESULT 11

US-08-487-826B-13/c
Sequence 13, Application US/08487826B

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487, 826B
FILING DATE: 10-SEP-1993

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 7.1%; Score 51; DB 2; Length 19124;
Best Local Similarity 45.8%; Pred. No. 0.065;
Matches 171; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY 9 AAGTTATTTTAAATTTCTATTAAACATCTCTCAAGCATTATTATCTATATCTC 68
Db |||||
18383 ATATTTTATTTATTAATAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTT 18324
QY 69 ACTGAATTTTAAAGAAATACATTAGTATTAGAAAACTAGAAAAAGATAATGCAGAT 128
Db |||||
18323 ATGTGAATTAATCTTATGTTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 18264
QY 129 AATTAACCTTACATGAAAAAGAAATTTATAACAAGACTGAGAACGTTAATAATTGAA 188
Db |||||
18263 AGTTAATAATAAATAAGAAATTAAGTTATATTAATAAATAACAAGAAATTAATTGTAT 18204
QY 189 ATGAGATTATTAATTGAAAACTGCATCTGAAACCAACTTTATTGTTCAATTATNCTTAA 248
Db |||||
18203 ATTATATTATTTATTTTAAATAGGAAACCTATTAATATTGTATTATTAATAATTTTATA 18144
QY 249 TGATGTTGTTTATGACTAATACACTGATTTTCAAGAGAAACCAGTTAAAAATAT 308
Db |||||
18143 TGAGATTATATATTTTTTCGCTCGGATTATTGAGAAATTAGAGTAAACAAAAATAAAC 18084
QY 309 TTTTATTTTAAAAATAAGCCTGTGTTCAAGCTGTGATCATATTTCTTTATTTTGGATTG 368
Db |||||
18083 ACATATATAAACAATATATTAATACACATTTAAATAATATATTATAAATAATACAAA 18024
QY 369 GGAANAAAAATACT 381
Db |||||
18023 TAAATAAATAAAT 18011

RESULT 12

US-07-991-867B-23
Sequence 23, Application US/07991867B

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 678 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-07-991-867B-23

Query Match 7.0%; Score 50.2; DB 1; Length 678;
Best Local Similarity 46.1%; Pred. No. 0.069;
Matches 166; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 5 AGAAAAAGTATTATTTCTATTAACATCTCTCAAGCATTATTTCCTATA 64
DB 31 ATAATAATTTATTTGTTGTTATAGTATTATTAACATTTGTTAATATA 90
QY 65 TCTCACTGAATTTAAGAAATAACATAGATTAGAAAACTAGAAAAAGATAATGC 124
DB 91 GATTTTAAATAATAATAATATATATATATATATCAATTAACGCACTAATAAAC 150
QY 125 AGATAATTAACCTACATGAAAAAGGAAATTTATACAAAGGACTGAGAACGTTAAAT 184
DB 151 AATATAATTAATTTAAATTTATAGCATTATTCAGATATTAATTTTGACAAATTTTAAAC 210
QY 185 TGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAACTTATTGTTCAATTATNC 244
DB 211 AATAATAATTAATCTTTAGTAAACACAGCTAATTAATTTACAAAGATATACCAATATTTAAT 270
QY 245 TTAATGATGTTTATGACTAATACACTGATTTTTCAGAGAAGAAACCACTTAAAT 304
DB 271 GTAATAATATATATCTTAATCAATATATTTTTCAGCGCTAGTAATATGTAAT 330
QY 305 ATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGA 364
DB 331 ATATTATTAGATTAAAGAAAAACATTAATATATAAGAAATCCATTTTATTTTGA 390

RESULT 13
US-08-107-755A-23
Sequence 23, Application US/08107755A
Patent No. 5721352
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville

STATE: Florida
COUNTRY: U.S.A.
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107,755A
FILING DATE: 19-AUG-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,658
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 678 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-107-755A-23

Query Match 7.0%; Score 50.2; DB 1; Length 678;
Best Local Similarity 46.1%; Pred. No. 0.069;
Matches 166; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 5 AGAAAAAGTATTATTTCTATTAACATCTCTCAAGCATTATTTCCTATA 64
DB 31 ATAATAATTTATTTGTTGTTATAGTATTATTAACATTTGTTAATATA 90
QY 65 TCTCACTGAATTTAAGAAATAACATAGATTAGAAAACTAGAAAAAGATAATGC 124
DB 91 GATTTTAAATAATAATAATATATATATATATATCAATTAACGCACTAATAAAC 150
QY 125 AGATAATTAACCTACATGAAAAAGGAAATTTATACAAAGGACTGAGAACGTTAAT 184
DB 151 AATAATAATTAATCTTTAGTAAACACAGCTAATTAATTTTTCAGAAATTTTAAAC 210
QY 185 TGAATGAGATTATTAATTTGAAAACTGCATCTGAAAGCAACTTATTGTTCAATTATNC 244
DB 211 ATAATAATATATCTTTAGTAAACACAGCTAATAATTTTTCAGAGATATACCAATATTTAAT 270
QY 245 TTAATGATGTTTATGACTAATACACTGATTTTTCAGAGAAGAAACCACTTAAAT 304
DB 271 GTAATAATATATCTTAATCAATATATTTTTCAGCGCTAGTAATATGTAAT 330
QY 305 ATATTTTATTTTAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGA 364
DB 331 ATATTATTAGATTAAAGAAAAACATTAATATATAAGAAATCCATTTTATTTTGA 390

RESULT 14
US-08-544-332-23
Sequence 23, Application US/08544332
Patent No. 5935777
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 678 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-544-332-23

Query Match      7.0%; Score 50.2; DB 2; Length 678;
Best Local Similarity 46.1%; Pred. No. 0.069;
Matches 166; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 5 AGAAAAAGTATTCTAATTTCTATTAAACATCTCTCAAGCATTATTTCCTATA 64
   |||||
Db 31 ATAAATATTATTGTAATTAGTATTATTAATTAACATTGCTTAATATA 90
QY 65 TCTCACTGAATTTAAGAAATAACATTAGTATTAGAAAACTAGAAAAAAGATAATGC 124
   |||||
Db 91 GATTTTAAATAATAATAATAATATATATATATCATATTAACGCACTAATATAAC 150
QY 125 AGATAATTAACTTACATGAAGAAAGAAATTAACAAGAGCTGAGAACGTTAAAT 184
   |||||
Db 151 AATATAATAATTAAATTATACGATTATTACAGATATTATATTTTGACAAATTTAAC 210
QY 185 TGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTATTGTTCAATTATNC 244
   |||||
Db 211 ATAAATAATAATCTTTAGTAACACAAGCTAATTAATTAACAAGATATACCAATATTAAAT 270
QY 245 TTAATGATGTGTTTATGACTAATACACTGATTTTCAAGAAAGCAACCAATGTTAAAA 304
   |||||
Db 271 GTAAATAATAATATATCTAATCAATATAATTTTATTCAGCGCTAGTAATAATGTTAAAT 330
QY 305 ATATTTTATTTTAAAAAAGAGCGTGTGTTCAAGCTCTGATCATATTTCTTTAATTTTGA 364
   |||||
```

```

Db 331 ATATTATTAGATTAGAAAAACATTAAATATAAGAAATCCATTTTATATTGTA 390

RESULT 15
US-09-370-861A-23
Sequence 23, Application US/09370861A
Patent No. 6410221
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 6410221el Entomopoxvirus Expression System
FILE REFERENCE: UF114.C4.D1
CURRENT APPLICATION NUMBER: US/09/370,861A
CURRENT FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: US 07/991,867
PRIOR FILING DATE: 1992-12-07
PRIOR APPLICATION NUMBER: US 08/107,755
PRIOR FILING DATE: 1993-08-19
PRIOR APPLICATION NUMBER: WO 92/14818
PRIOR FILING DATE: 1992-02-12
PRIOR APPLICATION NUMBER: US 07/827,685
PRIOR FILING DATE: 1992-01-30
PRIOR APPLICATION NUMBER: US 07/657,584
PRIOR FILING DATE: 1991-02-19
SOFTWARE: PatentIn version 3.1
SEQ ID NO 23
LENGTH: 678
TYPE: DNA
ORGANISM: Amsacta moorei entomopoxvirus
US-09-370-861A-23

Query Match      7.0%; Score 50.2; DB 4; Length 678;
Best Local Similarity 46.1%; Pred. No. 0.069;
Matches 166; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

QY 5 AGAAAAAGTATTCTAATTTCTATTAAACATCTCTCAAGCATTATTTCCTATA 64
   |||||
Db 31 ATAAATATTATTGTAATTAGTATTATTAATTAACATTGCTTAATATA 90
QY 65 TCTCACTGAATTTAAGAAATAACATTAGTATTAGAAAACTAGAAAAAAGATAATGC 124
   |||||
Db 91 GATTTTAAATAATAATAATAATATATATATATCATATTAACGCACTAATATAAC 150
QY 125 AGATAATTAACTTACATGAAGAAAGAAATTAACAAGAGCTGAGAACGTTAAAT 184
   |||||
Db 151 AATATAATAATTAAATTATACGATTATTACAGATATTATATTTTGACAAATTTAAC 210
QY 185 TGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTATTGTTCAATTATNC 244
   |||||
Db 211 ATAAATAATAATCTTTAGTAACACAAGCTAATTAATTTTCAAGATATACCAATATTAAAT 270
QY 245 TTAATGATGTGTTTATGACTAATACACTGATTTTCAAGAAAGCAACCAATGTTAAAA 304
   |||||
Db 271 GTAAATAATAATATCTAATCAATATAATTTTATTCAGCGCTAGTAATAATGTTAAAT 330
QY 305 ATATTTTATTTTAAAAAATAGCGTGTGTTCAAGCTCTGATCATATTTCTTTAATTTTGA 364
   |||||
Db 331 ATATTATTAGATTAGAAAAACATTAAATATAAGAAATCCATTTTATATTGTA 390
   |||||
```

Search completed: November 27, 2003, 12:32:16
Job time : 53.701 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:02:42 ; Search time 2782.58 Seconds
(without alignments)
10100.295 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 687

Sequence: 1 attctaattctcttataaac.....cttctgtatgcacagtcgtc 687

Scoring table: IDENTITY_NUC

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srb:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_srb:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrc:*
38: em_ey:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	683	99.4	687	6	ARI146580	ARI146580 Sequence
2	683	99.4	687	6	BD079829	BD079829 Cancer-as
3	638.4	92.9	1052	9	HSSCPXT16	U11313 Human stero
4	638.4	92.9	2572	9	HUMSCP2A	M75883 Human stero
5	638.4	92.9	122176	2	AL358233	AL358233 Homo stero
6	638.4	92.9	175046	2	AC022728	AC022728 Homo sapi
7	638.4	92.9	193774	9	AL445183	AL445183 Human DNA
8	636.8	92.7	1439	9	BC005911	BC005911 Homo sapi
9	611.8	89.1	714	6	ARI146579	ARI146579 Sequence
10	611.8	89.1	714	6	BD079828	BD079828 Cancer-as
11	455.6	66.3	1229	9	S52450	S52450 sterol carr
12	440.4	64.1	1219	9	HUMSTEA	M55421 Human stero
13	411.8	59.9	843	6	ARI146582	ARI146582 Sequence
14	411.8	59.9	843	6	BD079831	BD079831 Cancer-as
15	296.4	43.1	1500	9	HUMSCP2B	M75884 Human stero
16	216.4	31.5	2661	4	AF051897	AF051897 Oryctolag
17	180	26.2	145564	2	AC134794	AC134794 Mus muscu
18	179	26.1	2626	10	BC018384	BC018384 Mus muscu
19	179	26.1	248677	2	AL844206	AL844206 Mus muscu
20	177.4	25.8	2599	10	RAT60KDA	M62763 Rat 60 kDa
21	168.8	24.6	2571	10	RATSCPXA	M57453 Rat sterol
22	161	23.4	173	6	AR246194	AR246194 Sequence
23	154.6	22.5	904	10	S80339	S80339 SCP-2=stero
24	126.2	18.4	263	6	A74403	A74403 Sequence 89
25	126.2	18.4	263	6	A77382	A77382 Sequence 89
26	85.8	12.5	2152	6	AX306091	AX306091 Sequence
27	85.8	12.5	2152	10	MUSSCP	M91458 Mus musculu
28	81.8	11.9	95477	9	AC007076	AC007076 Homo sapi
29	77	11.2	168799	9	AC009531	AC009531 Homo sapi
30	75.8	11.0	125623	3	AC115599	AC115599 Dictyoste
31	74.8	10.9	113202	9	HSJ397H23	AL121972 Human DNA
32	74.6	10.9	147956	2	AC137839	AC137839 Medicago
33	74.4	10.8	152878	3	CEY18D10A	AL034393 Caenorhab
34	73.8	10.7	149597	2	AC034271	AC034271 Homo sapi
35	73.8	10.7	157749	2	AC025535	AC025535 Homo sapi
36	73.2	10.7	251762	3	AE014851	AE014851 Plasmodi
37	73.2	10.7	310779	2	AC005140	AC005140 Plasmodi
38	72.6	10.6	1192	9	HSA323759	AJ323759 Homo sapi
39	72.6	10.6	158548	3	PFMAL3P2	AL034558 Plasmodi
40	72.2	10.5	5850	3	DDIDDP2	M55298 Dictyosteli
41	72.2	10.5	182870	3	AC116960	AC116960 Dictyoste
42	72	10.5	151259	2	AC138822	AC138822 Homo sapi
43	72	10.5	160701	9	AC108106	AC108106 Homo sapi
44	72	10.5	172404	9	AC138948	AC138948 Homo sapi
45	72	10.5	181154	2	AC138842	AC138842 Homo sapi

ALIGNMENTS

RESULT 1
ARI146580
LOCUS ARI146580 687 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 20 from patent US 6218521.
ACCESSION ARI146580
VERSION ARI146580.1 GI:15109769
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 687)
AUTHORS Obata,Y.
TITLE Isolated nucleic acid molecules associated with gastric cancer and methods for diagnosing and treating gastric cancer
JOURNAL Patent: US 6218521-A 20 17-APR-2001;

Pred. No. 16 the number of results predicted by chance to have a

FEATURES		Location/Qualifiers			
source		1. .687 /organism="unknown"			
BASE COUNT	242 a	98 c	85 g	258 t	4 others
ORIGIN					
Query Match					
99.4%; Score 683; DB 6; Length 687;					
Best Local Similarity 100.0%; Pred. No. 1.4e-91;					
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATTTAATTTCTAATTAACATTTCTCTCAAGCATTTTATCCTATATCTCACTGAA	60		
Db	1	ATTTAATTTCTAATTAACATTTCTCTCAAGCATTTTATCCTATATCTCACTGAA	60		
QY	61	TTTTAANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA	120		
Db	61	TTTTAANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA	120		
QY	121	ACTTACATGAAAAAGAAAAATTATAACAAGAGCTGAGAACGTTATAAATTGAAATGAGA	180		
Db	121	ACTTACATGAAAAAGAAAAATTATAACAAGAGCTGAGAACGTTATAAATTGAAATGAGA	180		
QY	181	TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATATTCTTAATGATGG	240		
Db	181	TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATATTCTTAATGATGG	240		
QY	241	TGTTTATGACTAATACACTGATTTTTCATATAAGGAAACCCATGTTAAAAATATTTTAT	300		
Db	241	TGTTTATGACTAATACACTGATTTTTCATATAAGGAAACCCATGTTAAAAATATTTTAT	300		
QY	301	TTTAAAAATAAGCCTGTGTCTCAAGCTCTGATCATATTTCTTTATTGTTGGGAAGA	360		
Db	301	TTTAAAAATAAGCCTGTGTCTCAAGCTCTGATCATATTTCTTTATTGTTGGGAAGA	360		
QY	361	AAATACTGTTTCGATAGCATGAAATGCCAAAAATTTTAGATTTTAACTCNCCTAATTTT	420		
Db	361	AAATACTGTTTCGATAGCATGAAATGCCAAAAATTTTAGATTTTAACTCNCCTAATTTT	420		
QY	421	AAGAACTATTGAGAAATTGATTATGACATGAAAGTGCACAACACTAATTACTGCCAGCT	480		
Db	421	AAGAACTATTGAGAAATTGATTATGACATGAAAGTGCACAACACTAATTACTGCCAGCT	480		
QY	481	GTTGCATGTGTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC	540		
Db	481	GTTGCATGTGTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC	540		
QY	541	NGAATAACCTTAATATACTTTGTTAGCCAAACAAACTTTTGTGTTACATAGTCTTT	600		
Db	541	NGAATAACCTTAATATACTTTGTTAGCCAAACAAACTTTTGTGTTACATAGTCTTT	600		
QY	601	GGATTTTACTGTTCTTAATTTTATTTCTGAAACTCCATTTTCCCGAGACCAATAATTACC	660		
Db	601	GGATTTTACTGTTCTTAATTTTATTTCTGAAACTCCATTTTCCCGAGACCAATAATTACC	660		
QY	661	TATTTAATTTGTTATGCACAGTTGTT	687		
Db	661	TATTTAATTTGTTATGCACAGTTGTT	687		
RESULT 2					
LOCUS	BD079829	687 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Cancer-associated nucleic acids and polypeptides.				
ACCESSION	BD079829				
VERSION	BD079829.1	GI:22625432			
KEYWORDS	JP 2001516009-A/495.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 687)				
AUTHORS	Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I., Oghare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.				

TITLE	Cancer-associated nucleic acids and polypeptides									
JOURNAL	Patent: JP 2001516009-A 495 25-SEP-2001;									
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH									
OS	Homo sapiens (human)									
PN	JP 2001516009-A/495									
PD	25-SEP-2001									
PF	15-JUL-1998 JP 2000503425									
PR	17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR									
10-OCT-1997 US	60/061765,10-OCT-1997 US 08/948705 PR									
11-OCT-1997 GB	9721697.2,22-JUN-1998 US 09/102322 PI LLOYD									
J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI	TSENG									
CHEN,										
PI	IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI									
OZLEM TURECI,										
PI	UGUR SAHIN									
PC	G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC									
A61P35/00,										
PC	C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,									
PC	C12N15/00									
CC	Cancer-associated nucleic acids and polypeptides. FH Key									
Location/Qualifiers										
FT	source 1..687									
FT	Location/Qualifiers /organism='Homo sapiens (human)'									
FEATURES										
source	1..687									
/organism="Homo sapiens"										
/mol_type="genomic DNA"										
/db_xref="taxon:9606"										
BASE COUNT	242 a 98 c 85 g 258 t 4 others									
ORIGIN										
Query Match	99.4%; Score 683; DB 6; Length 687;									
Best Local Similarity	100.0%; Pred. No. 1.4e-91;									
Matches 687; Conservative	0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATTTAATTTCTATTAAACATCTTCTCAAGCATTTATCTCTATATCTCACTGAA	60							
Db	1	ATTTAATTTCTATTAAACATCTTCTCAAGCATTTATCTCTATATCTCACTGAA	60							
QY	61	TTTTAANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA	120							
Db	61	TTTTAANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA	120							
QY	121	ACTTACATGAAAAAGGAAATTATACAAAGAGCTGAGAACGTTATAATTGAAATGAGA	180							
Db	121	ACTTACATGAAAAAGGAAATTATACAAAGAGCTGAGAACGTTATAATTGAAATGAGA	180							
QY	181	TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATATTCTTAATGATGG	240							
Db	181	TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATATTCTTAATGATGG	240							
QY	241	TGTTTATGACTAATACACTGATTTTTCATAATAAGGAAACCCATGTTAAAAATATTTTAT	300							
Db	241	TGTTTATGACTAATACACTGATTTTTCATAATAAGGAAACCCATGTTAAAAATATTTTAT	300							
QY	301	TTTAAAAATAAGCCTGTGTCTCAAGCTCTGATCATATTTCTTTATTGATTGGGAAGA	360							
Db	301	TTTAAAAATAAGCCTGTGTCTCAAGCTCTGATCATATTTCTTTATTGATTGGGAAGA	360							
QY	361	AAATACTGTTTCTGATAGCATGAAATGCAAAAATTTTAGATTTTAACTCNCCTAATTTT	420							
Db	361	AAATACTGTTTCTGATAGCATGAAATGCAAAAATTTTAGATTTTAACTCNCCTAATTTT	420							
QY	421	AAGAACTATTGAGAAATTGATTATGACATGAAAGTGCACAACACTAATTACTGCCAGCT	480							
Db	421	AAGAACTATTGAGAAATTGATTATGACATGAAAGTGCACAACACTAATTACTGCCAGCT	480							
QY	481	GTTGCATGTGTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTAACTGAACTTTCAGC	540							
Db	481	GTTGCATGTGTTCTTACTTAGTCTCTCCCAAGGAAAACTCTTAACTGAACTTTCAGC	540							
QY	541	NGAATAACCTTAATATACTTTGTTAGCCAAACAAACTTTTGTGTTACATAGTCTTT	600							

|||||
Db 541 NGATAACCTTAATATACTTTGTTAGCCAAACAACCTTTTGTACATAGTCTTT 600
Qy 601 GCATTTACTGTTCTTAATTTATCTGAACCTCCATTTTCCCGACGACCATTAATTAACC 660
Db 601 GGATTTTACTGTTCTTAATTTATCTGAACCTCCATTTTCCCGACGACCATTAATTAACC 660
Qy 661 TATTTAATTTGTTATGACAGAGTTGTT 687
Db 661 TATTTAATTTGTTATGACAGAGTTGTT 687
RESULT 3
HSSCPX16/c
LOCUS HSSCPX16 1052 bp DNA linear PRI 25-JAN-1997
DEFINITION Human sterol carrier protein-X/sterol carrier protein-2
(SCP-X/SCP-2) gene, exon 16, and complete cds.
ACCESSION U11313
VERSION U11313.1 GI:532077
KEYWORDS 16 of 16
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1052)
AUTHORS Ohba,T., Rennett,H., Pfeifer,S.M., He,Z., Yamamoto,R., Holt,J.A.,
Billheimer,J.T. and Straus,J.F. III.
TITLE The structure of the human sterol carrier protein X/sterol carrier
protein 2 gene (SCP2)
JOURNAL Genomics 24 (2), 370-374 (1994)
MEDLINE 95213031
PUBMED 7698762
REFERENCE 2 (bases 1 to 1052)
AUTHORS Straus,J.F. III.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1994) Jerome F. Straus III, Department of
Obstetrics and Gynecology, Division of Reproductive Biology,
University of Pennsylvania Medical Center, 778 Clinical Research
Building, 422 Curie Boulevard, Philadelphia, PA 19104-6142, USA
FEATURES
source location/Qualifiers
1.1052
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p32"
/cbase_type="liver"
join(U11297.1:1..3666,U11299.1:1..108,U11300.1:1..122,
U11301.1:1..182,U11302.1:1..115,U11303.1:1..177,
U11304.1:1..114,U11305.1:1..137,U11306.1:1..201,
U11307.1:1..198,U11308.1:1..158,U11309.1:1..1725,
U11310.1:1..153,U11311.1:1..180,U11312.1:1..130,1..1034)
/partial
/gene="SCP-X/SCP-2"
join(U11297.1:3452..3641,U11299.1:26..83,U11300.1:26..97,
U11301.1:26..157,U11302.1:26..90,U11303.1:26..152,
U11304.1:26..89,U11305.1:26..112,U11306.1:26..176,
U11307.1:26..173,U11308.1:26..133,U11309.1:1547..1700,
U11310.1:26..128,U11311.1:26..155,U11312.1:26..105,
26..1034)
/gene="SCP-X/SCP-2"
/product="sterol carrier protein-X/sterol carrier
protein-2"
join(U11297.1:3573..3641,U11299.1:26..83,U11300.1:26..97,
U11301.1:26..157,U11302.1:26..90,U11303.1:26..152,
U11304.1:26..89,U11305.1:26..112,U11306.1:26..176,
U11307.1:26..173,U11308.1:26..133,U11309.1:1547..1700,
U11310.1:26..128,U11311.1:26..155,U11312.1:26..105,
26..121)
/gene="SCP-X/SCP-2"
/codon_start=1
/product="sterol carrier protein-X/sterol carrier

protein-2"
/protein_id="AAB41286.1"
/db_xref="GI:1773239"
/translation="MSSSPWEPAALRRVVGVMGKTFVKPGAEHSRDYDPLAEAGK
KALADAOIPYSAVDOACVGVVGDSTCGORAIYHSLGMDGIPINNNCATGSTALF
MARQLIQGVAAECVLALEPKRSKSLGKIPSDRTIPTDKVDLLINKYGLSAHPVAP
QMFYAGKEHMEKYGTKIEHFAKIGMKNHKSNNPVSQFQDEYSLDEVMASKEVDF
LTILOCCPTSDGAAAILASEAFVQKYGLOSKAVEILAQEMWTDLPSSFEEKSIKMW
GFMSKEAARKCYEKSGLTPNDIDVIELHDCFSSTNELTYEALGCPREGOGATLVDRG
DNTYGGKVINPSGGLISKHPDGLAQCAELCWOLRGEAGKQVPGAKVALQHNL
DIGAVVVTLYKMGFPPEAASSRTHQIEAVPTSSASDGFKANLVFKEIEKLEEGEO
FVKKIGIIFAFKVKGDPGKEATWVDVKNKGSVLPNSDKADCTITMADSDFLALM
TGKXNPQSAFFQGLKITGNMGLMKLQNLQLPGNAKL"
26..1034
/gene="SCP-X/SCP-2"
/note="nucleotides 122-1027 are not translated"
/number=16
BASE COUNT 352 a 157 c 164 g 379 t
ORIGIN
Query Match 92.9%; Score 638.4; DB 9; Length 1052;
Best Local Similarity 98.1%; Pred. No. 4.6e-85;
Matches 675; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
Qy 1 ATTTAATTTTCTATTAAACATCTTCCTCAAGCATTTATTCCTATATCTCACTGAA 60
Db 1014 ATTTAATTTTCTATTAAACATCTTCCTCAAGCATTTATTCCTATATCTCACTGAA 955
Qy 61 TTTAANAATAACATTAGTATTAGAAAACTAGGAAAAAGATNAATGAGATAATTAA 120
Db 954 TTTAAGAATAACATTAGTATTAGAAAACTAGGAAAAAGATAAATGAGATAATTAA 895
Qy 121 ACTTACATGAAAAAGAAATTATTAACAAGAGCTGAGAACGTTATAATTGAATGAGA 180
Db 894 ACTTACATGAAAAAGAAATTATTAACAAGAGAGCTGAGAACGTTATAATTGAATGAGA 835
Qy 181 TTATAATTTGAAAACTGCATCTGAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 240
Db 834 TTATAATTTGAAAACTGCATCTGAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 775
Qy 241 TGTTTATGACTAATACACTGATTTTCAATAAGGAAACCATGTTAAAAATATTTTAT 300
Db 774 TGTTTATGACTAATACACTGATTTTCAATAAGGAAACCATGTTAAAAATATTTTAT 715
Qy 301 TTTAAAAATTAAGCCTGTGTTCAAGCTTGATCATATTTCTTTATTTGATTTGGGAAGA 360
Db 714 TTTAAAAATTAAGCCTGTGTTCAAGCTTGATCATATTTCTTTATTTGATTTGGGAAGA 655
Qy 361 AAATACGTGTTCTGATAGCATGAAGTGAAGAAATTTTAGATTTTAACTGCTAATTTT 420
Db 654 AAATACGTGTTCTGATAGCATGAAGTGAAGAAATTTTAGATTTTAACTGCTAATTTT 595
Qy 421 AAGACTATTGAGAAATTTGATTAATGACATGAAGTGACACAACACTAACTGAGCCAGCT 480
Db 594 AAGACTATTGAGAAATTTGATTAATGACATGAAGTGACACAACACTAACTGAGCCAGCT 535
Qy 481 GTTGCAATTGCTTTCTTACTAGTTCCTCCAGAGAAACTCTTAACCTGAATCTTCAGC 540
Db 534 GTTGCAATTGCTTTCTTACTAGTTCCTCCAGAGAAACTCTTAACCTGAATCTTCAGC 475
Qy 541 NGAAATAA-CCTTAATATACTTTGTTAGCCAAACAATAA-CTTTTGTGTTTACATAGTTCT 598
Db 474 AGAATAATCTTTAATATACTTTGTTAGCCAAACAATAAAGCTTTTGTGTTTACATAGTTCT 415
Qy 599 TTGATTTTACTGTTCTTAATTTTATTTGAAAACCTCCATTTTCCCGACCATTAATTAC 658
Db 414 TTGATTTTACTGTTCTTAATTTTATTTGAAAACCTCAATTTTACCCCGACCATTAATTAC 355
Qy 659 CCTATTTAATCTTGTATGACAGAGTTGT 686
Db 354 CATA-TTAATCTTGTATGACAGAGTTGT 328

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
HUMSCP2A/LOCUS	Human sterol carrier protein X/sterol carrier protein 2 mRNA,	2572 bp	mRNA	linear	PRI 06-DEC-1993							
DEFINITION	Human sterol carrier protein X/sterol carrier protein 2 mRNA,											
ACCESSION	complete cds.											
VERSION												
KEYWORDS												
SOURCE												
ORGANISM												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												
MEDLINE												
PUBMED												
REFERENCE												
AUTHORS												
TITLE												
JOURNAL												

		/gene="SCP-X/SCP-2"	
		/evidence=experimental	
BASE COUNT	794 a 439 c 570 g 769 t		
ORIGIN			
Query Match	92.9%;	Score 638.4;	DB 9; Length 2572;
Best Local Similarity	98.1%;	Pred. No. 3.5e-85;	
Matches 675; Conservative	0;	Mismatches 10;	Indels 3; Gaps 3;
QY	1	ATTTTAATTTCTATTAACATCTCTCTCAAGCATTTTATCCTATATCTCACTGAA	60
Db	2558	ATTTTAATTTCTATTAACATCTCTCTCAAGCATTTTATCCTATATCTCACTGAA	24999
QY	61	TTTTAANAATAACATTAGTATTAGAAAAAAGTGAAGAAAAAGATNATGCAGATAATTAA	120
Db	2498	TTTTAAGAATAACATTAGTATTAGAAAAAAGTGAAGAAAAAGATNATGCAGATAATTAA	24399
QY	121	ACTTACATGAAAAAGGAAAAATTATAACAAAGGACTGAGAACGTTATAATTGAATTGAGA	180
Db	2438	ACTTACATGAAAAAGGAAAAATTATAACAAAGGACTGAGAACGTTATAATTGAATTGAGA	23799
QY	181	TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATTCTTAATGATGG	240
Db	2378	TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATTCTTAATGATGG	23199
QY	241	TGTTTATGACTAAATACCTGATTTTTTCAATTAAGGAAACCATGTTAAAAATATTTTAT	300
Db	2318	TGTTTATGACTAAATACCTGATTTTTTCAATTAAGGAAACCATGTTAAAAATATTTTAT	2259
QY	301	TTTAAATAAGCCTGTGTTCAGCTCTGATCATATTCTTTTATTTGATTTGGGAGAA	360
Db	2258	TTTAAATAAGCCTGTGTTCAGCTCTGATCATATTCTTTTATTTGATTTGGGAGAA	21999
QY	361	AAATACTGTTTCGATGACATGAAATGCAAAATTTTGAATTTTAACTCNCATAATTTT	420
Db	2198	AAATACTGTTTCGATGACATGAAATGCAAAATTTTGAATTTTAACTCNCATAATTTT	2139
QY	421	AAGAACTATTGAGAAATTGATTAATGACATGAAGTGCAACAACATAATTACTGGCCAGCT	480
Db	2138	AAGAACTATTGAGAAATTGATTAATGACATGAAGTGCAACAACATAATTACTGGCCAGCT	2079
QY	481	GTTGGCATGTGTCTCTTAAGTTAGTCTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC	540
Db	2078	GTTGGCATGTGTCTCTTAAGTTAGTCTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC	2019
QY	541	NGAATAA-CCTTAAATATACCTTTGTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT	598
Db	2018	AGAATAATCCTTAAATATACCTTTGTAGCCAAACAAAAGCTTTTGTGTTACATAGTTCT	19599
QY	599	TTGATTTTACTGTTCCTAATTTTATTTCTGAAACCTCAATTTTCCCGAGACCATAATTAC	658
Db	1958	TTGATTTTACTGTTCCTAATTTTATTTCTGAAACCTCAATTTTCCCGAGACCATAATTAC	1899
QY	659	CCTATTACTTTGTTATGCACAGTTGT 686	
Db	1898	CATA-TTAACCTTGTATAATGCACAGTTGT 1872	
RESULT 5			
LOCUS	AL358233	122176 bp	DNA linear HTG 10-JUL-2001
DEFINITION	Homo sapiens chromosome 1 clone RP5-835A17, *** SEQUENCING IN PROGRESS ***, 8 unordered pieces.		
ACCESSION	AL358233		
VERSION	AL358233.3	GI:9797852	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_CANCELLED.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	McLay, K.		
TITLE	1 Direct Submission		

JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 12, 2000 this sequence version replaced gi:9214208.

Center: Sanger Centre
Genome Center

Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk

Center project name: dj835A17
Project Information

Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator ET-amersham; 20% of reads Chemistry:

Dye-terminator Big Dye; 79% of reads

Consensus quality: 11880 bases at least Q40

Consensus quality: 120216 bases at least Q30

Consensus quality: 120979 bases at least Q20

Insert size: 121476; sum-of-contigs

Insert size: 136681; 7.7% error; agarose-fp

Quality coverage: 4.05x in Q20 bases; sum-of-contigs Quality coverage: 3.75x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 13486: contig of 13486 bp in length
* 13487 13586: gap of 100 bp
* 13587 30893: contig of 17307 bp in length
* 30894 30993: gap of 100 bp
* 30994 37022: contig of 6029 bp in length
* 37023 37122: gap of 100 bp
* 37123 45381: contig of 8259 bp in length
* 45382 45481: gap of 100 bp
* 45482 56788: contig of 11307 bp in length
* 56789 56888: gap of 100 bp
* 56889 84391: contig of 27503 bp in length
* 84392 84491: gap of 100 bp
* 84492 117843: contig of 33352 bp in length
* 117844 117943: gap of 100 bp
* 117944 122176: contig of 4233 bp in length.

FEATURES Location/Qualifiers

source

1. 122176

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clone_lib="RP5-835A17"

1. 13486

misc_feature

/note="assembly_fragment:01065
fragment_chain:1"

misc_feature

13587. 30893
/note="assembly_fragment:00181
fragment_chain:1"

misc_feature

30994. 37022
/note="assembly_fragment:00153
fragment_chain:1"

misc_feature

37123. 45381
/note="assembly_fragment:00580
fragment_chain:2"

misc_feature

45482. 56788
/note="assembly_fragment:00948
fragment_chain:2"

misc_feature

56889. 84391
/note="assembly_fragment:00192"

misc_feature

84492. 117843

misc_feature /note="assembly_fragment:00299.0"
117944. 122176
/note="assembly_fragment:00926"
BASE COUNT 31142 a 30481 c 29651 g 30201 t 701 others
ORIGIN

Query Match 92.9%; Score 638.4; DB 2; Length 122176;
Best Local Similarity 98.1%; Pred. No. 1e-85;
Matches 675; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

QY 1 ATTTAATTTTCTATTAAACATTTCTCAAGCATTTTATCTATATCTCACTGAA 60
DB 47294 ATTTAATTTTCTATTAAACATTTCTCTCAAGCATTTTATCTATATCTCACTGAA 47353
QY 61 TTTTAAANAATAACATTTAGTATGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 120
DB 47354 TTTTAAANAATAACATTTAGTATGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 47413
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGACTGAAACGTTATTAATTGAAATGAGA 180
DB 47414 ACTTACATGAAAAAGAAAAATTATACAAAGACTGAAACGTTATTAATTGAAATGAGA 47473
QY 181 TTATAATTTGAAAACCTGCATCTGAAGCAAACTTATTTGTTCAATTAATCTTAATGATGG 240
DB 47474 TTATAATTTGAAAACCTGCATCTGAAGCAAACTTATTTGTTCAATTAATCTTAATGATGG 47533
QY 241 TGTTTATGACTAATACACTGATTTTCAATTAAGAAACCATGTTAAAAATATTTTAT 300
DB 47534 TGTTTATGACTAATACACTGATTTTCAATTAAGAAACCATGTTAAAAATATTTTAT 47593
QY 301 TTTAAAAATAAGCCTGTGTTCAAGCTGATCATATTCTTTATTTTGGGAAGA 360
DB 47594 TTTAAAAATAAGCCTGTGTTCAAGCTGATCATATTCTTTATTTTGGGAAGA 47653
QY 361 AAATACGTTTCTGATAGCAAGAAATGCAAAATTTTATGATTTTAACTCNCATAATTTT 420
DB 47654 AAATACGTTTCTGATAGCAAGAAATGCAAAATTTTATGATTTTAACTCNCATAATTTT 47713
QY 421 AAGACATTTGAGAAATTGATTAATGACATGAAGTGACAAACATTAATTAATCTGGCCAGCT 480
DB 47714 AAGACATTTGAGAAATTGATTAATGACATGAAGTGACAAACATTAATTAATCTGGCCAGCT 47773
QY 481 GTTGCAATGTTGTTCTTACTTACTTACTTCTCCCAAGAAACCTCTTAACTGAACTCTCAGC 540
DB 47774 GTTGCAATGTTGTTCTTACTTACTTACTTCTCCCAAGAAACCTCTTAACTGAACTCTCAGC 47833
QY 541 NGAATAA-CCTTAATATATCTTTGTTAGCCAAACAAA-CCTTTTGTGTTACATAGTTCT 598
DB 47834 NGAATAATCCTTAATATATCTTTGTTAGCCAAACAAA-CCTTTTGTGTTACATAGTTCT 47893
QY 599 TTGATTTTACTGTTCTCTAATTTTATCTGAAACTCCATTTTCCCGAGACCATTAATTAC 658
DB 47894 TTGATTTTACTGTTCTCTAATTTTATCTGAAACTCCATTTTCCCGAGACCATTAATTAC 47953
QY 659 CCTATTTAATCTTGTATGACAGTTGT 686
DB 47954 CATA-TTAACTTTGTAATGACAGTTGT 47980

RESULT 6
AC022728/c 175046 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 1 clone RP11-310J14 map 1, WORKING DRAFT
DEFINITION SEQUENCE, 34 unordered pieces.
ACCESSION AC022728
VERSION AC022728.4 GI:7249198
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 175046)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 1, clone RP11-310J14
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175046)
AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavkiy,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A.,
Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczkyl,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Menus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE Direct Submission
JOURNAL Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 175046)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavkiy,L., Bouhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Chepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczkyl,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigillio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE Direct Submission
JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 16, 2000 this sequence version replaced gi:6980310.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5491

Center clone name: 310_J_14

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156135 bases at least Q40

Consensus quality: 165373 bases at least Q30

Consensus quality: 168980 bases at least Q20

Insert size: 18500; agarose-fp

Insert size: 171746; sum-of-ctigs

Quality coverage: 3.4 in Q20 bases; agarose-fp

Quality coverage: 3.6 in Q20 bases; sum-of-ctigs

----- NOTE: This is a 'working draft' sequence. It currently

* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 56: contig of 56 bp in length
57 156: gap of 100 bp
157 1191: contig of 1035 bp in length
1192 1291: gap of 100 bp
1292 2432: contig of 1141 bp in length
2433 2532: gap of 100 bp
2533 3742: contig of 1210 bp in length
3743 3842: gap of 100 bp
3843 4921: contig of 1079 bp in length
4922 5021: gap of 100 bp
5022 6073: contig of 1052 bp in length
6074 6173: gap of 100 bp
6174 7552: contig of 1379 bp in length
7553 7652: gap of 100 bp
7653 9001: contig of 1349 bp in length
9002 9101: gap of 100 bp
9102 10622: contig of 1521 bp in length
10623 10722: gap of 100 bp
10723 11836: contig of 1114 bp in length
11837 11936: gap of 100 bp
11937 13398: contig of 1462 bp in length
13399 13498: gap of 100 bp
13499 14949: contig of 1451 bp in length
14949 15049: gap of 100 bp
15050 16260: contig of 1211 bp in length
15050 16261: gap of 100 bp
16261 16360: contig of 1905 bp in length
16361 18265: contig of 1551 bp in length
18266 18365: gap of 100 bp
18366 19916: contig of 1551 bp in length
19917 20016: gap of 100 bp
20017 22939: contig of 2923 bp in length
22940 23039: gap of 100 bp
23040 26445: contig of 3406 bp in length
26446 26545: gap of 100 bp
26546 29327: contig of 2782 bp in length
29328 29427: gap of 100 bp
29428 32200: contig of 2773 bp in length
32201 32300: gap of 100 bp
32301 36738: contig of 4438 bp in length
36739 36838: gap of 100 bp
36839 43447: contig of 6609 bp in length
43448 43547: gap of 100 bp
43548 48525: contig of 4978 bp in length
48526 48625: gap of 100 bp
48626 52791: contig of 4166 bp in length
52792 52891: gap of 100 bp
52892 58391: contig of 5500 bp in length
58392 58491: gap of 100 bp
58492 64869: contig of 6378 bp in length
64870 64969: gap of 100 bp
64969 73369: contig of 8400 bp in length
73370 73469: gap of 100 bp
73470 81101: contig of 7632 bp in length
81102 81201: gap of 100 bp
81202 87754: contig of 6553 bp in length
87755 87854: gap of 100 bp
87855 95534: contig of 7680 bp in length
95535 95634: gap of 100 bp
95635 102797: contig of 7163 bp in length
102798 102897: gap of 100 bp
102898 111817: contig of 8920 bp in length
111818 111917: gap of 100 bp
111918 121160: contig of 9243 bp in length
121161 121260: gap of 100 bp
121261 140537: contig of 19277 bp in length
140538 140637: gap of 100 bp

```

* 140638 175046: contig of 34409 bp in length.
FEATURES
    source
        1. .175046
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /map="1"
            /clone="RP11-310J14"
            /clone_1lb="RPC1-11 Human Male BAC"
misc_feature
    1. .56
        /note="assembly_fragment"
        clone_end:T7
        vector_side:right"
misc_feature
    157. .1191
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    1292. .2432
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    2533. .3742
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    3843. .4921
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    5022. .6073
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    6174. .7552
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    7653. .9001
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    9102. .10622
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    10723. .11836
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    11937. .13398
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    13499. .14949
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    15050. .16260
        /note="assembly_fragment"
        clone_end:T7
misc_feature
    16361. .18265
        /note="assembly_fragment"
        clone_end:T7
Query Match
    92.9%; Score 638.4; DB 2; Length 175046;
Best Local Similarity 98.1%; Pred. No. 9.4e-86;
Matches 675; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
QY 1 ATTTAATTTCTATTAACATCTCTCTCAAGCATTTTATCCTATCTCACTGAA 60
    |||
DB 43344 ATTTAATTTCTATTAACATCTCTCTCAAGCATTTTATCCTATCTCACTGAA 43285
QY 61 TTTAANNAATACATTAGTATTAGAAAAACCTAGAAAAAGATNAATGCAGTAATTAA 120
    |||
DB 43284 TTTAAGAAATACATTAGTATTAGAAAAACCTAGAAAAAGATNAATGCAGTAATTAA 43225
QY 121 ACTTACATGAAAAAGAAATTTAACAAGAGCTGAGAACGTTATTAATGAAATGAGA 180
    |||
DB 43224 ACTTACATGAAAAAGAAATTTAACAAGAGCTGAGAACGTTATTAATGAAATGAGA 43165
QY 181 TTATATTTGAAACTGCATCTGAAAGCAACTTATTGTTCAATTATCTTAATGATGG 240
    |||
DB 43164 TTATATTTGAAACTGCATCTGAAAGCAACTTATTGTTCAATTATCTTAATGATGG 43105
QY 241 TGTATTGACTAATACACTGATTTTCAATAAGAAACCAGTAAATAATTTTAT 300
    |||
DB 43104 TGTATTGACTAATACACTGATTTTCAATAAGAAACCAGTAAATAATTTTAT 43045
QY 301 TTTAAAAATAAGCCTGTCTCAAGCTCTGATCATATTCTTTTATTGATTTGGGAAGA 360
    |||
DB 43044 TTTAAAAATAAGCCTGTCTCAAGCTCTGATCATATTCTTTTATTGATTTGGGAAGA 42985
QY 361 AAATACTGTTCTGATGACATGAAATGCAAAATTTTATGATTTTAAATCTCNCATATTT 420
    |||
```

```

DB 42984 AAATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAATCTCACTAATTT 42925
QY 421 AAGAACTATTGAGAAATTTGATTAATGACATGAGTGCACAACTAATTAATCTGCCAGCT 480
    |||
DB 42924 AAGAACTATTGAGAAATTTGATTAATGACATGAGTGCACAACTAATTAATCTGCCAGCT 42865
QY 481 GTTGCAATTGTGTTCTTACTAGTTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 540
    |||
DB 42864 GTTGCAATTGTGTTCTTACTAGTTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 42805
QY 541 NGAAATA-CCTTAATATCTTTGTTAGCCAAACAAA-CCTTTTGTGTTACATAGTTCT 598
    |||
DB 42804 AGAATTAATCCTTAATATCTTTGTTAGCCAAACAAA-CCTTTTGTGTTACATAGTTCT 42745
QY 599 TTGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCCATTTTCCCAAGACCATTAATTAC 658
    |||
DB 42744 TTGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTTACCCAGACCATTAATTAC 42685
QY 659 CCTAATTTACTGTTCTTATGACAGTTGT 686
    |||
DB 42684 CATA-TTAACTTTGTAATGACAGTTGT 42658
RESULT 7
AL445183/C 193774 bp DNA linear PRI 05-APR-2002
LOCUS Human DNA sequence from clone RP11-334A14 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL445183
VERSION AL445183.19 GI:20068427
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
    1 Wallis, J.
        Direct Submission
        Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
        Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
        humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
        On Apr 7, 2002 this sequence version replaced gi:17939714.
        During sequence assembly data is compared from overlapping clones.
        where differences are found these are annotated as variations
        together with a note of the overlapping clone name. Note that the
        variation annotation may not be found in the sequence submission
        corresponding to the overlapping clone, as we submit sequences with
        only a small overlap as described above.
        This sequence was finished as follows unless otherwise noted: all
        regions were either double-stranded or sequenced with an alternate
        chemistry or covered by high quality data (l.e., phred quality >=
        30); an attempt was made to resolve all sequencing problems, such
        as compressions and repeats; all regions were covered by at least
        one plasmid subclone or more than one M13 subclone; and the
        assembly was confirmed by restriction digest. The following
        abbreviations are used to associate primary accession numbers given
        in the feature table with their source databases: Em:, EMBL; Sw:,
        SWISSPROT; Tr:, TrEMBL; Wp:, WormPEP; Information on the WormPEP
        database can be found at
        http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
        was generated from part of bacterial clone contigs of human
        chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
        Group. Further information can be found at
        http://www.sanger.ac.uk/HGP/Chr1
        RP11-334A14 is from the library RPC1-11.2 constructed by the group
        of Pieter de Jong. For further details see
        http://www.chori.org/bacpac/home.htm
        VECTOR: pBACe3.6.
FEATURES
    source
        1. .193774
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
```


QY	361	AAATACTGTTTCTGATAGCATGAAATGC AAAATTTT	TAGATTTT	TAAATCTC	NC	TAAATTTT	420
Db	1038	AAATACTGTTTCTGATAGCATGAAATGC AAAATTTT	TAGATTTT	TAAATCTC	NC	TAAATTTT	979
QY	421	AAGAACTATTGAGAAATGATTTAATGACATGAAGTGC	ACACACTAATTA	CTGGCCAGCT			480
Db	978	AAGAACTATTGAGAAATGATTTAATGACATGAAGTGC	ACACACTAATTA	CTGGCCAGCT			919
QY	481	GTTGGCATTGTGTTTCTTAAGTTCTCCCAAGGAAAA	CTTAAACTGAATCTT	CAGC			540
Db	918	GTTGGCATTGTGTTTCTTAAGTTCTCCCAAGGAAAA	CTTAAACTGAATCTT	CAGC			859
QY	541	NGAATAA-CCTTAATAATACCTTGTAGCCAAACAAA	-CTTTTGTGTTA	CATAGTTCT			598
Db	858	AGAATAATCCTTAATAATACCTTGTAGCCAAACAAA	AGCTTTTGTGTTA	CATAGTTCT			799
QY	599	TTGGATTTTACTGTTCCCTAATTTTATTTCTGAAC	TCCCATTTTCCC	GAGACCATTA	TTAC		658
Db	798	TTGGATTTTACTGTTCCCTAATTTTATTTCTGAAC	TCCCATTTTCCC	GAGACCATTA	TTAC		739
QY	659	CCTATTTAATCTTGTGTTATGCACAGTTGT	686				
Db	738	CATA-TTAACCTTGTGAATGCACAGTTGT	712				

RESULT	9
ARI146579	
LOCUS	ARI146579 714 bp DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 19 from patent US 6218521.
ACCESSION	ARI146579
VERSION	ARI146579.1 GI:15109768
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 714)
TITLE	Obata,Y.
JOURNAL	Isolated nucleic acid molecules associated with gastric cancer and methods for diagnosing and treating gastric cancer
FEATURES	Patent: US 6218521-A 19 17-APR-2001; Location/Qualifiers 1..714 source /organism="unknown"
BASE COUNT	258 a 100 c 92 g 260 t 4 others
ORIGIN	

Query Match	89.1%;	Score 611.8;	DB 6;	Length 714;
Best Local Similarity	97.1%;	Pred. No. 4.3e-81;		
Matches 671;	Conservative	0;	Mismatches 15;	Indels 5;
			Gaps	5;
QY	1	ATTTTAATTTTCTATTAAACATCTCTCTCAAAAGCATTATTTTATCTCTATATCTCACTGAA	60	
DB	15	ATTTTAATTTTCTATTAAACATCTCTCTCAAAAGCATTATTTTATCTCTATATCTCACTGAA	74	
QY	61	TTTTAANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA	120	
DB	75	TTTTAAGAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA	134	
QY	121	ACTTACATGAAAAAGAAATTTAACAAGAAGCTGAGACGTTATTAATTGAATGAGA	180	
DB	135	ACTTACATGAAAAAGAAATTTAACAAGAAGCTGAGACGTTATTAATTGAATGAGA	194	
QY	181	TTATTAATTTGAAAACCTGCATCTGAAAGCAAACTTATTTGTTCAATTATCTTAATGATGG	240	
DB	195	TTATTAATTTGAAAACCTGCATCTGAAAGCAAACTTATTTGTTCAATTATCTTAATGATGG	254	
QY	241	TGTTTATGACTAATACACTGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT	300	
DB	255	TGTTTATGACTAATACACTGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT	314	
QY	301	TTTAAAAATAAGCCTGTGTTCAAGCTGATCATATTTCTTTATTTTGAATTTGGGAAGA	360	
DB	315	TTTAAAAATAAGCCTGTGTTCAAGCTGATCATATTTCTTTATTTTGAATTTGGGAANA	374	

[illegible]

RESULT	10
BD079828	
LOCUS	BD079828
DEFINITION	Cancer-associated nucleic acids and polypeptides.
ACCESSION	BD079828
VERSION	BD079828.1 GI:22625431
KEYWORDS	JP 2001516009-A/494.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 714) Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I., Oghare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U. Cancer-associated nucleic acids and polypeptides Patent: JP 2001516009-A 494 25-SEP-2001; LUDWIG INSTITUTE FOR CANCER RESEARCH OS Homo sapiens (human) PN JP 2001516009-A/494 PD 25-SEP-2001 PR 15-JUL-1998 JP 2000503425 PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR 10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR 11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG CHEN, PI IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCUH,P OZLEM TURECİ, PI UGUR SAHİN
TITLE	JOURNAL
COMMENT	
FEATURES	source source
BASE COUNT	258 a 100 c 92 g 260 t 4 others
ORIGIN	
Query Match	89.1%; Score 611.8; DB 6; Length 714;

Best Local Similarity 97.1%; Pred. No. 4.3e-81;
Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

```
QY 1 ATTTAATTTCTATTAACATCTTCTCAAGCATTTATCTTATCTCAGTAA 60
    |||||
Db 15 ATTTAATTTCTATTAACATCTTCTCAAGCATTTATCTTATCTCAGTAA 74
QY 61 TTTAANAATAACATAGTATTAGAAAACTAGGAAAAAGATNAATGCAGATATTAA 120
    |||||
Db 75 TTTTAAGAAATACATAGTATTAGAAAACTAGGAAAAAGATNAATGCAGATATTAA 134
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGGAGCTGAGAACGTTATAATTGAAATGAGA 180
    |||||
Db 135 ACTTACATGAAAAAGAAAAATTATACAAAGGAGCTGAGAACGTTATAATTGAAATGAGA 194
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATCTTAATGATGG 240
    |||||
Db 195 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATCTTAATGATGG 254
QY 241 TGTTTATGACTAATACATGATTTTTCATAGGAAACCATGTTAAAAATATTTTAT 300
    |||||
Db 255 TGTTTATGACTAATACATGATTTTTCATAGGAAACCATGTTAAAAATATTTTAT 314
QY 301 TTTAAAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTTGATTTGGAGAGA 360
    |||||
Db 315 TTTAAAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTTGATTTGGAGANA 374
QY 361 AAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCATAATTTT 420
    |||||
Db 375 AAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCATAATTTT 434
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGCAACAACATAATTAAGTGGCCAGCT 480
    |||||
Db 435 AANAACATTGAGAAATGATTAATGACATGAGTGCAACAACATAATTAAGTGGCCAGCT 494
QY 481 GTTGCAATGTGTTCTTACTTAGTTCTCCCAA--GGAAAACCTCTAAACGTGAATCTTCAG 539
    |||||
Db 495 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAAGGAAAACTCTTAATGAACTTTCAG 554
QY 540 CNGAATAA--CCTTAATATATCTTTGTTAGCCAAACAAA--CTTTTGTGTTACATAGTTC 597
    |||||
Db 555 CAGATAATCCTTAATATATCTTTGTTAGCAAAACAAAAGCTTTTGTGTTACATAGTTC 614
QY 598 TTT-GGATTTTACTGTTCTCTAATTTTATTCTGAAACTCCATTTTCCCGACCATTAATT 656
    |||||
Db 615 TTTGGATTTTACTGTTCTCTAATTTTATTCTGAAACTCAATTTTACCCGACCATTAATT 674
QY 657 ACCCTATTTAACCTTGTATGACACAGTTGTT 687
    |||||
Db 675 ACCATA-TTAACTTTGTTNGACACAGTTGTT 704
```

RESULT 11

LOCUS S52450/c 1229 bp mRNA linear PRI 08-MAY-1993
DEFINITION sterol carrier protein 2 [human, liver, mRNA, 1229 nt].
ACCESSION S52450
VERSION S52450.1 GI:263550

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1229)
AUTHORS Yamamoto, R.
TITLE Localization of human sterol carrier protein 2 gene and cDNA
expression in COS-7 cell

JOURNAL Hokkaido Igaku Zasshi 67 (6), 839-848 (1992)
MEDLINE 93131254
PUBMED 1483685

GenBank staff at the National Library of Medicine created this
entry [NCBI gibs9 122504] from the original journal article.
This sequence comes from Fig. 1.

FEATURES Map location: 1.
Source Location/Qualifiers
1..1229
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

gene 1..1229
/gene="sterol carrier protein 2, SCP2"
84..515
/gene="sterol carrier protein 2, SCP2"
/note="this sequence comes from Fig. 1; SCP2"

CDS /codon_start=1
/product="sterol carrier protein 2"
/protein_id="AAB24921.1"
/db_xref="GI:263551"

/translation="MGFPEAASSFRTHQIEAVPTSSASDGFKANLVFEIEKLEEBG
EQFVKRIGGFIAFKVKDGPGEKEDTVVDVKNQGSVLPNSDKKADCTITMAASDFLA
LMTGRNPDQSAFFQGPLKITGNMGLAMKLQNLQDPNAXL"

BASE COUNT 409 a 200 c 255 g 365 t
ORIGIN

Query Match 66.3%; Score 455.6; DB 9; Length 1229;
Best Local Similarity 97.2%; Pred. No. 3.7e-58;
Matches 494; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

```
QY 181 TTATAATTTGAAAACCTGCATCTGAAAGCAAACTTTATTTGTTCAATTAATCTAATGATGG 240
    |||||
Db 1228 TTTTATTTTGAAGAACTGCATCTGAAAGCAAACTTTATTTGTTCAATTAATCTAATGATGG 1169
QY 241 TGTTTATGACTAATACACTGATTTTTCATAGGAAACCCATGTTAAAAATATTTTAT 300
    |||||
Db 1168 TGTTTATGACTAATACACTGATTTTTCATAGGAAACCCATGTTAAAAATATTTTAT 1109
QY 301 TTTAAAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTTTGAATTTGGAGAGA 360
    |||||
Db 1108 TTTAAAAATAAGCCGTGTCTCAAGCTCTGATCATATTTCTTTATTTTGAATTTGGAGAGA 1049
QY 361 AAATACTGTTCTGATAGCATGAATGCAAAAATTTTAGATTTTAACTCNCATAATTTT 420
    |||||
Db 1048 AAATACTGTTCTGATAGCATGAATGCAAAAATTTTAGATTTTAACTCNCATAATTTT 989
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTAAGTGGCCAGCT 480
    |||||
Db 988 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTAAGTGGCCAGCT 929
QY 481 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAGGAAACCTCTTAACCTGAATCTTCAGC 540
    |||||
Db 928 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAGGAAACCTCTTAACCTGAATCTTCAGC 869
QY 541 NGAATAA--CCTTAATATATCTTTGTTAGCCAAACAAA--CTTTTGTGTTACATAGTTC 598
    |||||
Db 868 AGAATAATCCTTAATATATCTTTGTTAGCAAAACAAAAGCTTTTGTGTTACATAGTTC 809
QY 599 TTGATTTTACTGTTCTCTAATTTTATCTGAAACTCCATTTTCCCGACCATTAATTAC 658
    |||||
Db 808 TTGATTTTACTGTTCTCTAATTTTATCTGAAACTCAATTTTACCCGACCATTAATTAC 749
QY 659 CCTATTTAACCTTGTATGACACAGTTGT 686
    |||||
Db 748 CATA-TTAACTTTGTTAATGACACAGTTGT 722
```

RESULT 12

LOCUS HUMSTEA/c 1219 bp mRNA linear PRI 06-DEC-1993
DEFINITION Human sterol carrier protein-2 (SCP-2) mRNA, complete cds.
ACCESSION M55421
VERSION M55421.1 GI:432978

KEYWORDS sterol carrier protein-2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE	1 (bases 1 to 1219)
AUTHORS	Yamamoto,R., Kallen,C.B., Babalola,G.O., Rennert,H., Billheimer,J.T. and Straus,J.F. III.
TITLE	Cloning and expression of a cDNA encoding human sterol carrier protein 2
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 88 (2), 463-467 (1991)
MEDLINE	91110550
PUBMED	1703300
REFERENCE	2 (bases 1 to 1219)
AUTHORS	Vesa,J., Hellsten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T., Mead,S., Cowell,J.K., Straus,J.F.III. and Peltonen,L.
TITLE	Assignment of sterol carrier protein X/sterol carrier protein 2 to 1p32 and exclusion as the causative gene for infantile neuronal ceroid lipofuscinosis
JOURNAL	Unpublished
COMMENT	On Dec 6, 1993 this sequence version replaced gi:432972. Original source text: Human female liver.
FEATURES	Location/Qualifiers
source	1. 1219
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/map="1p32"
	/tissue_type="liver"
gene	1. 1219
	/gene="SCP-2"
CDS	84. 515
	/gene="SCP-2"
	/codon_start=1
	/product="sterol carrier protein-2"
	/protein_id="AAA03559.1"
	/db_xref="GI:432973"
	/translation="MGPEAASSFRHQIEAVPTSSASDGFKANLVEKEIEKLEEEG EQPVKKIGIFAFKVKDGPGEKATVVVDVKNKGSLPNSDKKADCTITMADSDFLA LMTGKNPQSAFFQGLKITGNMGLMKLQNLQPGNAKL"
sig_peptide	84. 143
	/gene="SCP-2"
mat_peptide	144. 512
	/gene="SCP-2"
	/product="unnamed"
polya_signal	782. 787
	/gene="SCP-2"
	/evidence=not_experimental
polya_signal	1062. 1067
	/gene="SCP-2"
	/evidence=not_experimental
polya_signal	1107. 1112
	/gene="SCP-2"
	/evidence=not_experimental
polya_signal	1190. 1195
	/gene="SCP-2"
	/evidence=experimental
polya_site	1219
	/gene="SCP-2"
	/evidence=experimental
BASE COUNT	402 a 195 c 255 g 367 t
ORIGIN	
Query Match	64.1%; Score 440.4; DB 9; Length 1219;
Best Local Similarity	97.6%; Pred. No. 6.5e-56;
Matches	488; Conservative 0; Mismatches 8; Indels 4; Gaps 4;
OY	189 TGAAGCTGCATCTGAAGCAAACTTTATGTTCAATTATCTTAATGATGGTTTTAT 248
Db	1219 TGAAGCTGCATCTGAAGCAAACTTTATGTTCAATTATCTTAATGATGGTTTTAT 1160
OY	249 GACTAATACACTGATTTTCAATAAGAAACCATGTTAAAAATATTTTATTTAAAAA 308
Db	1159 GACTAATACACTGATTTTCAATAAGAAACCATGTTAAAAATATTTTATTTAAAAA 1100
OY	309 TAAAGCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGAGAATAACTG 368
Db	1099 TAAAGCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGAGAATAACTG 1040

OY	369 TTTCTGATACATGAATGCAGAAATTTTATGATTTTATCTCNCNTAATTTTAAGACTA 428
Db	1039 TTTCTGATACATGAATGCAGAAATTTTATGATTTT-ATCTCACTAATTTTAAGACTA 981
OY	429 TTGAGAAATGATTAATGACATGAAGTCACAACACTAATTACTGGCAGCTGTGGCAT 488
Db	980 TTGAGAAATGATTAATGACATGAAGTCACAACACTAATTACTGGCAGCTGTGGCAT 921
OY	489 TGTGTTCTTACTTAGTCTCCCAAGGAAACTCTTAACTGAATCTTCAGCNGAATAA- 547
Db	920 TGTGTTCTTACTTAGTCTCCCAAGGAAACTCTTAACTGAATCTTCAGCAGATAAT 861
OY	548 CCTTAATAATACTTTGTTAGCCAAACAAG-CTTTTGTGTTACATAGTTCCTTGATTT 606
Db	860 CCTTAATAATACTTTGTTAGCCAAACAAGCTTTTGTGTTACATAGTTCCTTGATTT 801
OY	607 TACTGTTCTTAATTTTATTTCTGAAGCTCCATTTTCCCGAGACCATTAATACCCTATT 666
Db	800 TACTGTTCTTAATTTTATTTCTGAAGCTCAATTTTACCAGACCATTAATACCATA-TTA 742
OY	667 ACTTGTGTTATGCAGTGTG 686
Db	741 ACTTGTGTTATGCAGTGTG 722
RESULT 13	
ARI46582	843 bp DNA linear PAT 08-AUG-2001
LOCUS	ARI46582
DEFINITION	Sequence 22 from patent US 6218521.
ACCESSION	ARI46582
VERSION	ARI46582.1 GI:15109771
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	1 (bases 1 to 843)
AUTHORS	Obata,Y.
TITLE	Isolated nucleic acid molecules associated with gastric cancer and methods for diagnosing and treating gastric cancer
JOURNAL	Patent: US 6218521-A 22 17-Apr-2001;
FEATURES	Location/Qualifiers
source	1. 843
	/organism="unknown"
BASE COUNT	237 a 184 c 89 g 215 t 118 others
ORIGIN	
Query Match	59.9%; Score 411.8; DB 6; Length 843;
Best Local Similarity	71.8%; Pred. No. 1.2e-51;
Matches	495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;
OY	1 ATTTAATTTTCTATTAACATCTCTCTCAAGCATTTATTTATCTTATCTCACTGAA 60
Db	15 ATTTAATTTTCTATTAANCNTCTCCNCAANCAATTATTNACCCTATNNCNCNGAN 74
OY	61 TTTTAAATAATACATTAAGTATTAGAAAACTAGAAAAAAGATNAATGAGATAATTAA 120
Db	75 TTTNANAAANTACCTTNTNTNTTAAAAAACCTNGAAAAAATAATNGCAATANTTAA 134
OY	121 ACTTACATGAAGAAAGAAATTTATAACAAGGACTGAGACGTTATAATTGAATGAGA 180
Db	135 CCTTNCCTGAAGAAAGAAATTTNTACCAANGACNGAANCNTTNTAATNGAANTNAAA 194
OY	181 TTATAATTTGAAGAACTGCATCTGAAGCAAACTTATTTGTTCAATTATCTTAATGATGG 240
Db	195 TTATANTTNGAAANGCGCNCNGAAGAACCAANCTTNAATGCTCAATTATCTTNAANGAGGG 254
OY	241 TGTTTATGACTTAATACACTGATTTTTCATTAAGGAAACCATGTTAAAAATATTTTAT 300
Db	255 NNTTNAANNACTTAATNCCNGATTTTCCAATANGAANCCNNNTTAAANNTNTTNTAT 314
OY	301 TTTAAAAATAAGCTGTGTTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGGAAGA 360

Db 315 TTTAAAAATAACCCNGTNTCCAACCCNGATCANATTCCTTNATTGGATTGGGAAAA 374
QY 361 AAATACTGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTATCTCNCTAATTTT 420
Db 375 AAATNCNGTTCNNATACCNNGAANNCAANTTTTAAATTTTAAACCCCTANTTTT 434
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGCACACACTAATTACTGGCCAGCT 480
Db 435 AAANCTATNGAAAAANTNGATTANNGACTGGAATTGC-CAACCTANTTNCNGGCCACCN 493
QY 481 GTTGGCATTTGTTCTTCTTACTAGTTCTTCCCAAGGAAAACTTTAACTGAATCTTCAGC 540
Db 494 GTGGGCNTNGTNTCCCTTACTTANTCCCCCAAGGAAANNCCTTAANCNGAANTCCNCC 553
QY 541 NGAAATAACCTTAATATATCTTTGTTAGCCAAAC--AAAACTTTTGTGTTACATAGTTCT 598
Db 554 AAAATAACCCCTTAANTATCTTGGTAAACCAANCAAAACCTTTTNGTTTACNTANTCCT 613
QY 599 TTGGATTTTACTGTTCTCTAATTTTATTTCTGAACCTCCATTTTCCCGACAGACATAATTAC 658
Db 614 TGGGATTTAAACGGGTCCCAATTTNATCCNGAACCCANTTTTCCCCNAACCATANTTAC 673
QY 659 CCTATTTAATCTTGTGTTATGACAGTTGTT 687
Db 674 CAT-TTTACCTTGGTAAGCNCAGTNGTT 701

RESULT 14

BD079831 843 bp DNA linear PAT 27-AUG-2002
LOCUS BD079831
DEFINITION Cancer-associated nucleic acids and polypeptides.
ACCESSION BD079831
VERSION BD079831.1 GI:22625434
KEYWORDS JP 2001516009-A/497.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I.,
Oghare,J., Obata,Y., Pfeundschnub,M., Tureci,O. and Sahin,U.

TITLE Cancer-associated nucleic acids and polypeptides
JOURNAL Patent: JP 2001516009-A 497 25-SEP-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH

COMMENT OS Homo sapiens (human)
PN JP 2001516009-A/497
PD 25-SEP-2001
PF 15-JUL-1998 JP 2000503425

PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR
10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR
11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD
J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG
CHEN,

PI IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI
OZLEM TURECI,
PI UGUR SAHIN

PC G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC
A61P35/00,
PC C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
PC C12N15/00
CC Cancer-associated nucleic acids and polypeptides. FH Key
Location/Qualifiers
FT source 1..843
FT /Organism='Homo sapiens (human)'.
Location/Qualifiers
1..843
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
BASE COUNT 237 a 184 c 89 g 215 t 118 others
ORIGIN

Query Match 59.9%; Score 411.8; DB 6; Length 843;
Best Local Similarity 71.8%; Pred. No. 1.2e-51;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

QY 1 ATTTTATTTTCTATTTAAACATTTCTTCAAGCATTTATTTATCTCTACTGAA 60
Db 15 ATTTAAATTTCTTATTAANCNTCTCCNCAANCAATTATTNACCCTTATNCCNCGAN 74
QY 61 TTTTAAATAATACATTAAGTATTAGAAAACTAGGAAAAAGATNAATGAGATAATTAA 120
Db 75 TTTNAAAAANTACCTTNTNTTAAAAAACCTNGAAAAAAAATTAATNGCAATANTTAA 134
QY 121 ACTTACATGAAAAAGAAAAATTATTAACAAGACTGAGAACGTTATTAATGAAATGAGA 180
Db 135 CCTTNCCTGAAGAAAGAAATTTNTTACCAANGAANGAANCNTTNTTAATTNGAANTNAAA 194
QY 181 TTATATTTGAAAACTGCATCTGAAGCAAACTTATTTGTTCAATTAATCTTAATGATGG 240
Db 195 TTATANTTNGAAANCGCNCNGAAGCAACTTNATGTTCCATTAATCTTAANGAGGG 254
QY 241 TGTTTATGACTTAATACACTGATTTTTCATAAGGAAACCATGTAATAATTTTAT 300
Db 255 NNTTNAANACTTAATNCCNGATTTTCCAATANGAANCCNNNTTAAANTNTTTNAT 314
QY 301 TTTAAAAATAGCCCTGTGTTCAAGCTTGATCATATTTCTTTATTTGATTGGGAAGA 360
Db 315 TTTAAAAATAACCCNGTNTCCAACCCNGATCANATTCCTTNATTGGATTGGGAAAA 374
QY 361 AAATACTGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTATCTCNCTAATTTT 420
Db 375 AAATNCNGTTCNNATACCNNGAANNCAANTTTTAAATTTTAAACCCCTANTTTT 434
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGCACAACTAATTAATGAGCCAGCT 480
Db 435 AAANCTATNGAAAAANTNGATTANNGACTGGAATTGC-CAACCTANTTNCNGGCCACCN 493
QY 481 GTTGGCATTTGTTCTTCTTACTAGTTCTCCCAAGGAAAACTTTAACTGAATCTTCAGC 540
Db 494 GTGGGCNTNGTNTCCCTTACTTANTCCCCCAAGGAAANNCCTTAANCNGAANTCCNCC 553
QY 541 NGAAATAACCTTAATATATCTTTGTTAGCCAAAC--AAACCTTTTGTGTTACATAGTTCT 598
Db 554 AAAATAACCCCTTAANTATCTTGGTAAACCAANCAAAACCTTTTNGTTTACNTANTCCT 613
QY 599 TTGGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCCATTTTCCCGACAGACATAATTAC 658
Db 614 TGGGATTTAAACGGGTCCCAATTTNATCCNGAACCCANTTTTCCCCNAACCATANTTAC 673
QY 659 CCTATTTAATCTTGTGTTATGACAGTTGTT 687
Db 674 CAT-TTTACCTTGGTAAGCNCAGTNGTT 701

RESULT 15

HUMSCP2B/c 1500 bp mRNA linear PRI 06-DEC-1993
LOCUS HUMSCP2B
DEFINITION Human sterol carrier protein 2 mRNA, complete cds.
ACCESSION M75884
VERSION M75884.1 GI:432976
KEYWORDS sterol carrier protein-2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1500)

AUTHORS He,Z., Yamamoto,R., Futh,E.E., Schantz,L.J., Naylor,S.L.,
George,H., Billheimer,J.T. and Straus,J.F. III.

TITLE cDNAs encoding members of a family of proteins related to human
sterol carrier protein 2 and assignment of the gene to human
chromosome 1 p21---pter

JOURNAL DNA Cell Biol. 10 (8), 559-569 (1991)
MEDLINE 92029618
PUBMED 1718316

REFERENCE 2 (bases 1 to 1500)
AUTHORS Vesa,J., Hellesten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T.,
Mead,S., Cowell,J.K., Strauss,J.F.III, and Peltonen,L.
TITLE Assignment of sterol carrier protein X/sterol carrier protein 2 to
1p32 and exclusion as the causative gene for infantile neuronal
ceroid lipofuscionosis
JOURNAL Unpublished
COMMENT On Dec 6, 1993 this sequence version replaced gi:337996.
Original source text: Human liver cDNA to mRNA.
FEATURES
source
1. .1500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="1p32"
/feature_type="liver"
/clone_1b="Clontech HL-1001b,HL-1115b"
1. .1500
/gene="SCP-2"
82. .951
/gene="SCP-2"
/codon_start=1
/product="sterol carrier protein-2"
/protein_id="AA03558.1"
/db_xref="GI:432977"
/translation="MTDLPSFEEKSIIMVGFDMSEKAAKCYEKSGLTPNDIDVI
ELHDCFSNELLTYEALGLCPGGQATLVDRGDNITYGKQVYNPSGLISKHPLGAT
GLAOCABLWCQLRGEAGKROVPGAKVLAQNLGIGAVVVTLYKGFPEAASSFRTHQ
IEAVPTSSASDGFKANLVFKEIEKLEEGEQFVKKIGIFAFKVKYKDPGGEATWV
DVKNKGSLVLPNSDKADCTITMADSDFLALMTGKMPQSAFQGLKITGMGLAMK
LQNLQLPGNAKL"
polyA_signal
1218. .1223
/gene="SCP-2"
/note="no polyA or polyA signal was found at the end of
this cDNA clone, indicating that the mRNA sequence may be
partial at the 3' end"
/evidence=not experimental
BASE COUNT 482 a 254 c 342 g 422 t
ORIGIN
Query Match 43.1%; Score 296.4; DB 9; Length 1500;
Best Local Similarity 96.8%; Pred. No. 9.9e-35;
Matches 333; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 345 TTTGATTGGGAAGAAATACTGTTTCTGATGACATGAAATGCAAAATTTTAGATT 404
DB 1500 TTTGATTGGGAAGAAATACTGTTTCTGATGACATGAAATGCAAAATTTTAGATT 1441
QY 405 TAACTCTCNCCTAATTTTAAGAACTATTGAGAAATTGATTATGACATGAGTGCACAACAC 464
DB 1440 TAACTCTCNCCTAATTTTAAGAACTATTGAGAAATTGATTATGACATGAGTGCACAACAC 1381
QY 465 TAATTACTGGCCAGCTGTGGCATTGTGTTCTTACTTAGTTCTCCCAAGAAAACCTCTT 524
DB 1380 TAATTACTGGCCAGCTGTGGCATTGTGTTCTTACTTAGTTCTCTCCCAAGAAAACCTCTT 1321
QY 525 AAACCTGAATCTTCACGNGAATAA-CCTTAATAATACTTTGTTAGCCAAACAATAA-CTTTT 582
DB 1320 AAACCTGAATCTTCACGNGAATAA-CCTTAATAATACTTTGTTAGCCAAACAATAA-CTTTT 1261
QY 583 TTGTTACATAGTCTTTTGATTTTACTGTTCTTAATTTTATTTCTGAAACTCCATTTTTC 642
DB 1260 TTGTTACATAGTCTTTTGATTTTACTGTTCTTAATTTTATTTCTGAAACTCCATTTTTAC 1201
QY 643 CCCAGACCATAATTACCCTATTTAACCTTTGTTATGCACAGTTGT 686
DB 1200 CCCAGACCATAATTACCATA-TTAACTTTGTTAATGCACAGTTGT 1158

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:55:53 ; Search time 1851.59 Seconds
(without alignments)
9017.749 Million cell updates/sec

Title: US-09-835-992A-20
Perfect score: 687
Sequence: 1 attttaatttctctatctaac.....cttctgtaacgacagctgtc 687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	639	93.0	769	12	BQ014192 UI-H-ED1-BQ014192
2	638.4	92.9	761	12	BM997078 UI-H-ED0-BM997078
3	637.6	92.8	837	9	A1566109 UI-H-ED0-A1566109
4	636.8	92.7	738	12	BQ045161 UI-CF-EN1-BQ045161

5	635.8	92.5	712	13	B0619082	B0619082 UI-H-FH1-BQ008197
6	635.8	92.5	751	12	BQ008197	BQ008197 UI-H-ED1-B0619112
7	635.8	92.5	759	13	B0619112	B0619112 UI-H-FH1-A1675901
8	635.4	92.5	714	9	A1675901	A1675901 wC06h04.x
9	635.2	92.5	756	12	BM985376	BM985376 UI-CF-EC1-BM985376
10	635.2	92.5	762	13	B0627147	B0627147 UI-H-FG0-A1826287
11	634.4	92.3	770	9	A1826287	A1826287 wK33f07.x
12	632	92.0	721	12	BM968746	BM968746 UI-CF-DU1-CA424156
13	632	92.0	734	14	CA424156	CA424156 UI-H-FE1-AW052045
14	632	92.0	786	9	AW052045	AW052045 wx25e05.x
15	630.2	91.7	750	12	BQ009270	BQ009270 UI-H-ED1-BQ009270
16	629.8	91.7	734	13	B0933572	B0933572 AGENCOURT-A1565988
17	629.4	91.6	793	9	A1565988	A1565988 tn52a05.x
18	624	90.8	953	9	AL578584	AL578584 AL578584
19	623	90.7	672	12	BQ016778	BQ016778 UI-H-D10-BX414370
20	623	90.7	905	13	BX414370	BX414370 BX392565
21	617.8	89.9	901	13	BX392565	BX392565 BX392565
22	616.6	89.8	822	9	AV729461	AV729461 AV729461
23	616	89.7	821	9	A1640146	A1640146 wa29f11.x
24	609.4	88.7	876	9	A1831751	A1831751 wj40a10.x
25	606	88.2	671	9	A1683310	A1683310 tw50e02.x
26	605.4	88.1	784	12	B1861076	B1861076 603390807
27	600.8	87.5	699	10	BE739622	BE739622 601556314
28	599.2	87.2	824	10	BE738457	BE738457 601572489
29	592	86.2	728	9	A1831053	A1831053 wj62c05.x
30	591.4	86.1	1032	9	AL551280	AL551280 AL551280
31	590	85.9	840	9	AA659242	AA659242 nu12h08.s
32	589.4	85.8	648	12	BM311898	BM311898 i963e08.y
33	587.2	85.5	641	9	AW173415	AW173415 xj06d03.x
34	586.8	85.4	749	9	A1765940	A1765940 wh69d12.x
35	583.4	84.9	616	9	AW611897	AW611897 hg90b03.x
36	582.4	84.8	820	10	BF966918	BF966918 602286426
37	582.4	84.8	660	9	A1761686	A1761686 w162a07.x
38	577.6	84.1	673	9	AW512844	AW512844 xm04g12.x
39	570	83.0	702	9	A1431843	A1431843 ci25h04.x
40	569.6	82.9	637	13	BQ549593	BQ549593 ik88g10.x
41	569	82.8	727	9	A1636002	A1636002 wa14f07.x
42	568	82.7	744	10	BG569206	BG569206 602588480
43	566.4	82.4	627	9	A1224925	A1224925 q131c04.x
44	565.4	82.3	753	10	BE748610	BE748610 601571589
45	563.4	82.0	623	12	BM511625	BM511625 ijs1a05.x

ALIGNMENTS

RESULT 1
BQ014192 769 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-ED1-axs-g-24-0-UI.s1 NCI_CGAP_ED1 Homo sapiens CDNA clone
DEFINITION IMAGE:5833007 3', mRNA sequence.
ACCESSION BQ014192
VERSION BQ014192.1 GI:19739093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 769)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/RLNL at: http://image.llnl.gov
The following repetitive elements were found in this CDNA
sequence: 317-343, >AT_rich#low_complexity (matched complement)

Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. 769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5833007"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_LIB=UI-H-ED1
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"
BASE COUNT 272 a 110 c 100 g 285 t 2 others
ORIGIN
Query Match 93.0%; Score 639; DB 12; Length 769;
Best Local Similarity 98.1%; Pred. No. 3.7e-66;
Matches 675; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
QY 1 ATTTTAATTTCTATTAAACATCTTCTCAAGCATTATTTATCCTATATCTCACTGAA 60
Db 33 ATTTTAATTTCTATTAAACATCTTCTCAAGCATTATTTATCCTATATCTCACTGAA 92
QY 61 TTTTAAATAATACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 120
Db 93 TTTTAAATAATACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 152
QY 121 ACTTACATGAAAAAGGAAATTTATAACAAGGACTGAGAACGTTATNAATTGAATGAGA 180
Db 153 ACTTACATGAAAAAGGAAATTTATAACAAGGACTGAGAACGTTATNAATTGAATGAGA 212
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAATTATTGTTCAATTAATCTTAATGATGG 240
Db 213 TTATAATTGAAAACTGCATCTGAAAGCAAATTATTGTTCAATTAATCTTAATGATGG 272
QY 241 TGTTTTATGACTAATATACACTGATTTTTCATAAGGAAACCCAGTTAAAAATATTTTAT 300
Db 273 TGTTTTATGACTAATATACACTGATTTTTCATAAGGAAACCCAGTTAAAAATATTTTAT 332
QY 301 TTTTAAATAAGCCTGTGTTCAAGCTCGATCATATTCTTTATTTTGGGAAGA 360
Db 333 TTTTAAATAAGCCTGTGTTCAAGCTCGATCATATTCTTTATTTTGGGAAGA 392
QY 361 AAATACGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAATCTCNCCTAATTTT 420
Db 393 AAATACGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAATCTCNCCTAATTTT 452
QY 421 AAGAACTATTGAGAAATGATTAAATGATGAAGTGACACACACTAATTAATGCGCAGCT 480
Db 453 AAGAACTATTGAGAAATGATTAAATGATGAAGTGACACACACTAATTAATGCGCAGCT 512
QY 481 GTTGCAATGTGTTCTTCTTACTAGTTCTCCCAAGGAAAACTCTTAAACTGAATCTTCAGC 540
Db 513 GTTGCAATGTGTTCTTCTTACTAGTTCTCCCAAGGAAAACTCTTAAACTGAATCTTCAGC 572
QY 541 ANGAATAA-CCTTAATATACTTTGTTAGCCAAACAATA-CCTTTTGTGTTACATAGTTCT 598

|||||
Db 573 AGAATATCTCTTAATATATACCTTTGTAAGCAAAACAAAGCTTTTGTGTTACATAGTTCT 632
QY 599 TTGGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCCATTTTCCCGACGACATAATTAC 658
Db 633 TTGGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTTACCCGACGACATAATTAC 692
QY 659 CCTATTTAATCTTTGTTATGACAGTTGT 686
Db 693 CATA-TTAACTTTGNTATGCACAGTTGT 719
RESULT 2
BM997078 761 bp mRNA linear EST 26-MAR-2002
LOCUS UI-H-ED0-axo-f-03-0-UI.s1 NCI_CGAP_ED0 Homo sapiens cDNA clone
DEFINITION IMAGE:5831426 3', mRNA sequence.
ACCESSION BM997078 GI:19721979
VERSION BM997078.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 761)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
Sequence: 316-342, >At rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. 761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5831426"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED0"
/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI_CGAP_ED0 is a cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_LIB=UI-H-ED0
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"
BASE COUNT 269 a 108 c 98 g 285 t 1 others
ORIGIN
Query Match 92.9%; Score 638.4; DB 12; Length 761;
Best Local Similarity 98.1%; Pred. No. 4.4e-66;

Matches	675;	Conservative	0;	Mismatches	10;	Indels	3;	Gaps	3;
Qy	1	ATTTAATTTTCTATTAACA	TTCTTCTCAAGCATTA	TTTATCTATATCTCACTGAA	60				
Db	32	ATTTAATTTTCTATTAACA	TTCTTCTCAAGCATTA	TTTATCTATATCTCACTGAA	91				
Qy	61	TTTAAANAATAACATTAG	ATTAAGAAAACTAGAAAAA	GATNAATGACATAATTAA	120				
Db	92	TTTAAAGAATAACATTAG	ATTAAGAAAACTAGAAAAA	GATNAATGACATAATTAA	151				
Qy	121	ACTTACATGAAAAAGAAA	TTTAAACAAGGACTGAGA	CGTTATAATTTGAAATGACA	180				
Db	152	ACTTACATGAAAAAGAAA	TTTAAACAAGGACTGAGA	CGTTATAATTTGAAATGACA	211				
Qy	181	TTATAATTGMAAACTGCA	TCGMAAGCAAACTTATG	TTCATTAATTTCTTAATGATGG	240				
Db	212	TTATAATTGMAAACTGCA	TCGMAAGCAAACTTATG	TTCATTAATTTCTTAATGATGG	271				
Qy	241	TGTTTATGACTAATACAT	GATTTTTCATTAAGGAA	ACCAGTTAAATAATTTTAT	300				
Db	272	TGTTTATGACTAATACAT	GATTTTTCATTAAGGAA	ACCAGTTAAATAATTTTAT	331				
Qy	301	TTTAAATAAGCCTGTGT	CAAGCTCTGATCATATTT	CTTTATTTTGTGATTGGGA	360				
Db	332	TTTAAATAAGCCTGTGT	CAAGCTCTGATCATATTT	CTTTATTTTGTGATTGGGA	391				
Qy	361	AAATACGTCTTGATAGCA	TGAAATGCAAAATTTT	AGATTTTAAATCTCNC	420				
Db	392	AAATACGTCTTGATAGCA	TGAAATGCAAAATTTT	AGATTTTAAATCTCNC	451				
Qy	421	AAGAACTATTGAGAAAT	TGATTAATGACATGAAG	TGCACACACTAATTACTG	480				
Db	452	AAGAACTATTGAGAAAT	TGATTAATGACATGAAG	TGCACACACTAATTACTG	511				
Qy	481	GTTGGCATTGTGTTCTT	ACTTAAGTCTCCCAAGG	AAAACTCTTAAACTGAT	540				
Db	512	GTTGGCATTGTGTTCTT	ACTTAAGTCTCCCAAGG	AAAACTCTTAAACTGAT	571				
Qy	541	NGAATAA-CCTTAATATA	CTTTGTTAGCCCAACAAA	-CTTTTGTGTACATAGT	598				
Db	572	AGAATAATCCTTAATATA	CTTTGTTAGCCCAACAAA	-CTTTTGTGTACATAGT	631				
Qy	599	TTGATTTTACTGTTCC	TAAATTTTATTTCTGA	AACTCCATTTTCCCGA	658				
Db	632	TTGATTTTACTGTTCC	TAAATTTTATTTCTGA	AACTCAATTTTACCCG	691				
Qy	659	CCTATTTAACCTTGTAT	GCACAGTTGT 686						
Db	692	CATA-TTAACTTTGGT	ATGCAGTTGT 718						
RESULT 3									
LOCUS	AI566109	837 bp	mRNA	linear	EST 12-MAY-1999				
DEFINITION	tn53e08.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2172134 3'								
	similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR								
	(HUMAN);, mRNA sequence.								
ACCESSION	AI566109								
VERSION	AI566109.1	GI:4524561							
KEYWORDS	EST.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
REFERENCE	1 (bases 1 to 837)								
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .								
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),								
	Tumor Gene Index								
JOURNAL	Unpublished								
COMMENT	Contact: Robert Strausberg, Ph.D.								
	Email: cgapbs-r@mail.nih.gov								
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.								
	Emmert-Buck, M.D., Ph.D.								

FEATURES		Location/Qualifiers	
Source		1..837	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:2172134"	
		/lab_host="DH10B"	
		/clone_1fb="NCI_CGAP_Kid11"	
		/note="Organ: kidney; Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."	
BASE COUNT		300 a	129 c 118 g 285 t 5 others
ORIGIN			
Query Match	92.8%; Score 637.6; DB 9; Length 837;		
Best Local Similarity	97.5%; Pred. No. 5.2e-66;		
Matches 654; Conservative	0; Mismatches 16; Indels 1; Gaps 1;		
QY	1 ATTTTAATTTTCTAATTAACATTTCTCTCAAGCATTTATTTATCCTATATCTCACTGAA	60	
DB	18 ATTTTAATTTTCTAATTAACATTTCTCTCAAGCATTTATTTATCCTATATCTCACTGAA	77	
QY	61 TTTTANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGACATTAATTA	120	
DB	78 TTTTAAGAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGACATTAATTA	137	
QY	121 ACTTACATGAAAAAGAAAAATTATAACAAAGCATGAGAACGTTATAAATTGAAATGAGA	180	
DB	138 ACTTACATGAAAAAGAAAAATTATAACAAAGCATGAGAACGTTATAAATTGAAATGAGA	197	
QY	181 TTATAATTTGAAACTGCATCTGAAAGCAAACTTATTGTTCAATTATTCTTAATGATGG	240	
DB	198 TTATAATTTGAAACTGCATCTGAAAGCAAACTTATTGTTCAATTATTCTTAATGATGG	257	
QY	241 TGTTTATGACTAATACACTGATTTTTCATATAGGAAACCCATGTTAAAAATATTTTAT	300	
DB	258 TGTTTATGACTAATACACTGATTTTTCATATAGGAAACCCATGTTAAAAATATTTTAT	317	
QY	301 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGAGAGA	360	
DB	318 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGAGAGA	377	
QY	361 AAATACTGTTTCTGATAGCATGAAATGCAAAATTTTATGATTTTAAATCTCNCATAATTTT	420	
DB	378 AAATACTGTTTCTGATAGCATGAAATGCAAAATTTTATGATTTTAAATCTCNCATAATTTT	437	
QY	421 AAGAATCTATTGAGAAATTGATTTAATGACATGAAAGTGCACAACACTAATTACTGGCCAGCT	480	
DB	438 AAGAATCTATTGAGAAATTGATTTAATGACATGAAAGTGCACAACACTAATTACTGGCCAGCT	497	
QY	481 GTTGCAATGTGTCTTCTTACTTAGTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC	540	
DB	498 GTTGCAATGTGTCTTCTTACTTAGTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC	557	
QY	541 NGAATAA-CCTTAATATATACTTTGTTAGCCAAACAAACTTTTGTGTTACATAGTTCTT	599	

Db 558 AGAATAATCCTTAATATACTTGTGTAGCAACAAGCTNTNNTGTTTACATAGTCTCT 617

QY 600 TGGATTTTACTGTCTCCTAATTTTATTTCTGAAACTCCATTTTCCCGAGACCATTAATTACC 659

Db 618 TGGATTTTACTGTCTCCTAATTTTATTTCTGAAACTCAATTTTACCCGAGACCATTAATTACC 677

QY 660 CTATTTACTT 670

Db 678 ATATTTACTT 688

RESULT 4
BQ045161

LOCUS 738 bp mRNA linear EST 21-FEB-2003

DEFINITION UI-CF-EN1-aei-a-17-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone

ACCESSION BQ045161

VERSION BQ045161.1 GI:19796248

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo. 1 (bases 1 to 738)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).

The following repetitive elements were found in this cDNA sequence: 316-342, >AT rich#low_complexity (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

1. 738

location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-EN1-aei-a-17-0-UI"

/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-CF-EN1"

/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.

TAG LIB=UI-CF-EN1

TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr to LPS 24h

TAG SEQ=CTGCTCAGGT"

BASE COUNT 260 a 102 c 97 g 278 t 1 others

ORIGIN

Query Match 92.7%; Score 636.8; DB 12; Length 738;

Best Local Similarity 98.0%; Pred. No. 6.8e-66;

Matches 674; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 1 ATTTAATTTTCTATTAACATTTCTTCAAGCATTTTATCCTATATCTCACTGAA 60

Db 32 ATTTAATTTTCTATTAACATTTCTTCAAGCATTTTATCCTATATCTCACTGAA 91

QY 61 TTTAANAATAACATTAGATTAGAAAACTAGAAAAAGATNATGCAGATAATTAA 120

Db 92 TTTTAAAGAAATTAACATTAGATTAGAAAACTAGAAAAAGATTAATGCAGATAATTAA 151

QY 121 ACTTACATGAAAAAGAAATTTAACAAGAGCTGAGACGTTATAATTGAAATGACA 180

Db 152 ACTTACATGAAAAAGAAATTTAACAAGAGCTGAGACGTTATAATTGAAATGACA 211

QY 181 TTATATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATCTTAATGATGG 240

Db 212 TTATATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATCTTAATGATGG 271

QY 241 TGTTTATGACTAATACACGTATTTTCAATAAGAAACCATGTTAAAAATATTTTAT 300

Db 272 TGTTTATGACTAATACACGTATTTTCAATAAGAAACCATGTTAAAAATATTTTAT 331

QY 301 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGAGA 360

Db 332 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATTTGGAGA 391

QY 361 AATATCTGTTTCTGATAGCATGAATGCCAAATTTTATCTCNCCTAATTTT 420

Db 392 AATATCTGTTTCTGATAGCATGAATGCCAAATTTTATCTCNCCTAATTTT 451

QY 421 AAGAACTATTGAGAAATTGATTAATGACATGAAGTGACAACACTAATTACTGGCCAGCT 480

Db 452 AAGAACTATTGAGAAATTGATTAATGACATGAAGTGACAACACTAATTACTGGCCAGCT 511

QY 481 GTTGCATTGTGTTTCTTACTTAGTCTCTCCCAAGAAAACTTTAACTGAATCTTGACG 540

Db 512 GTTGCATTGTGTTTCTTACTTAGTCTCTCCCAAGAAAACTTTAACTGAATCTTGACG 571

QY 541 NGAAATAA-CCTTAATATACTTTGTTAGCCAAACAAAAA-CTTTTGTGTTACATAGTTCT 598

Db 572 AGAATAATCCTTAATATACTTTGTGAGCAAAACAAAGCTTTTGTGTTACATAGTTCT 631

QY 599 TTGGATTTTACTGTTCCCTAATTTTATTTCTGAAACTCCATTTTCCCGAGACCATTAATTAC 658

Db 632 TTGGATTTTACTGTTCCCTAATTTTATTTCTGAAACTCAATTTTACCAGACCATGATTAAC 691

QY 659 CCTATTTAATTGTTATGACAGATTGT 686

Db 692 CATTA-TTAACCTTGATGACAGATTGT 718

RESULT 5

LOCUS BU619082

DEFINITION UI-H-FH1-bfm-f-06-0-UI.s1 NCI CGAP_FH1 Homo sapiens cDNA clone

ACCESSION BU619082

VERSION BU619082.1 GI:23285297

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo. 1 (bases 1 to 712)

Oy		181	TTATAATTGGAAAACCTGCATCTGAAAGCAAACTTATTTGTTCATTAATTATCTTAATGATGG	240
Db		212	TTATAATTGAAAACCTGCATCTGAAAGCAAACTTATTTGTTCATTAATTATCTTAATGATGG	271
Oy	.	241	TGTTTTATGACTAATACACTGATTTTTTCAATAAGGAACCATGTTAAAAATATTTTAT	300
Db		272	TGTTTATGACTAATACACTGATTTTTTCAATAAGGAACCATGTTAAAAATATTTTAT	331
Oy		301	TTTAAAAATAGCCCTGTGTCCAAGCTCTGATCATATTTCTTTTATTTTGATTTGGAGA	360
Db		332	TTTAAAAATAGCCCTGTGTCCAAGCTCTGATCATATTTCTTTTATTTTGATTTGGAGA	391
Oy		361	AAATACTGTTTCTGATAGCAGAAATGCAAAATTTTGAATTTTAAATCTCNCTAATTTT	420
Db		392	AAATACTGTTTCTGATAGCAGAAATGCAAAATTTTGAATTTTAAATCTCACTAATTTT	451
Oy		421	AAGACTATTGAGAAATGATTAATGACATGAAGTGCACAACATAATTACTGGCCAGCT	480
Db		452	AAGACTATTGAGAAATGATTAATGACATGAAGTGCACAACATAATTACTGGCCAGCT	511
Oy		481	GTTGGCATTGTGTTTCTTAAGTTCTTCCCAGGAAAACCTTAAACTGAATCTTCAGC	540
Db		512	GTTGGCATTGTGTTTCTTAAGTTCTTCCCAGGAAAACCTTAAACTGAATCTTCAGC	571
Oy		541	NGAATAA-CCTTAATATATACTTTGTAGCCCAAACAANA--CTTTTGTGTACATAGTTC	597
Db		572	AGAATAATCCTTAATATATACTTTGTAGCCCAAACAANAAGCTTTTGTGTACATAGTTC	631
Oy		598	TTTGATTTTACTGTTCCCTAATTTTATTTCTGAAACTCCATTTTCCCGAGACCATAAATTA	657
Db		632	TTTGATTTTACTGTTCCCTAATTTTATTTCTGAAACTCAATTTTACCCCGAGACCATAAATTA	691
Oy		658	CCCTATTTAATCTTGTATATGACAGTTGT	686
Db		692	CCATA-TTAACCTTGTAAATGCAAGTTGT	719

```

RESULT 8
AI675901
LOCUS
DEFINITION
AI675901          714 bp      mRNA          linear      EST 17-DEC-1999
wc06h04.x1  NCI CGAP P-28 Homo sapiens cDNA clone IMAGE:214423 3'
similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR
(HUMAN);, mRNA sequence.
AI675901
AI675901          GI:4876381
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 714)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/bbrp/image/image.html
Insert length: 1429      Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 439.
location/Qualifiers
1. 714
/organism="Homo sapiens"
/mol_type="mRNA"

```

BASE COUNT	ORIGIN
257 a	98 c 92 g 263 t 4 others
	/db_xref="taxon:9606" /clone="IMAGE:2314423" /sex="male" /dev_stage="adult" /lab_host="DH10B" /clone_lib="NCI CGAP Pr28" /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

Query Match	92.5%	Score 635.4;	DB 9;	Length 714;
Best Local Similarity	97.1%;	Pred. No. 1e-65;		
Matches 664; Conservative	0;	Mismatches 18;	Indels 2;	Gaps .2;

QY		1	ATTTTAATTTTCTATTAAACATTCCTCTCAAGCATTATTTATCCATATCTCACTGAA	60
Db		12	ATTTTAATTTTCTATTAAACATTCCTCTCAAGCATTATTTATCCATATCTCACTGAA	71
QY		61	TTTTAANAATAACAATTAGTATTAGAAAAACTAGAAAAAAGATNAATGCAGATAATTAA	120
Db		72	TTTTAGAATAACATTAGTATTAGAAAAACTAGAAAAAAGATAAATGCAGATAATTAA	131
QY		121	ACTTACATGAAAAAGGAAAAATTATTAACAAAGGACTGAGAACGTATAAATGGAATGAGA	180
Db		132	ACTTACATGAAAAAGGAAAAATTATTAACAAAGGACTGAGAACGTATAAATGGAATGAGA	191
QY		181	TTATAATTGAAAACTGCATCTGAAAGCAAACCTTATTGTTCAATTATTCTTAATGATGG	240
Db		192	TTATAATTGAAAACTGCATCTGAAAGCAAACCTTATTGTTCAATTATTCTTAATGATGG	251
QY		241	TGTTTATGACTAATACATCATGATTTTTCATTAAGGAAACCATGTTAAAAATATTTTAT	300
Db		252	TGTTTATGACTAATACATCATGATTTTTCATTAAGGAAACCATGTTAAAAATATTTTAT	311
QY		301	TTTAAAATAAGCCGTGTTCACAGCTCGATCATATTTCTTTATTTTGATTGGGAAGA	360
Db		312	TTTAAAATAAGCCGTGTGTTCACAGCTCGATCATATTTCTTTATTTTGATTGGGAAGA	371
QY		361	AAATACTGTTCTGATAGCATGAAATGCCAAAATTTTAGATTTTAAATCTCNCTAATTTT	420
Db		372	AAATACTGTTCTGATAGCATGAAATGCCAAAATTTTAGATTTTAAATCTCACTAATTTT	431
QY		421	AAGAACTATTGAGAAATTGATTAATGACATGAAGTGACACAACATAATTACTGGCCAGCT	480
Db		432	AAGAACTATTGAGAAATTGATTAATGACATGAAGTGACACAACATAATTACTGGCCAGCT	491
QY		481	GTTGGCATTTGTCTTCTTACTTAGTCTCCCAAGGAAAACTGTAAACTGAATCTTCAGC	540
Db		492	GTTGGCATTTGTCTTCTTACTTAGTCTCCCAAGGAAAACTGTAAACTGAATCTTCAGC	551
QY		541	NGAATAA-CCTTAATATACTTTGTTAGCCAACAACAACTTTTGTTTACATAGTTCTT	599
Db		552	AGAATAATCCTTAATATATACITNCTAAGCAAAACAAAGCTTTTGTTTACATAGTTCTT	611
QY		600	TGGATTTTACTGTTCCCTAATTTTATTCTGAAACTCCATTTTTCGCCAGACCATAATTACC	659
Db		612	TGGATTTTACTGTTNCTAATTTTATTCTGAAACTCAATTTTACCACCAATTAATTACC	671
QY		660	CTAATTAACCTTGTATGCACAGT 683	
Db		672	ATA-TTAACTTTGTAATGCACAGT 694	

RESULT 9
BM985376

LOCUS BM985376 756 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-CF-EC1-acg-p-23-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
ACCESSION UI-CF-EC1-acg-p-23-0-UI 3', mRNA sequence.
VERSION BM985376
KEYWORDS BM985376.1 GI:19611803
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 756)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 316-342. >AT rich#Low_complexity (matched compliment)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source Location/Qualifiers
1..756

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acg-p-23-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGGCTTAC.
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGTGCTTAC"

BASE COUNT 268 a 107 c 97 g 283 t 1 others
ORIGIN

Query Match 92.5%; Score 635.2; DB 12; Length 756;
Best Local Similarity 97.8%; Pred. No. 1e-65;
Matches 673; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 1 ATTTAATTTCTATTAAACATCTCTCAAGCATATTATTTCTATATCTCACTGAA 60
J|||||

Db 32 ATTTAATTTCTATTAAACATCTCTCAAGCATATTATTTCTATATCTCACTGAA 91
QY 61 TTTTAAANAATAACATTTAGTATTAGAAAAAACTAGAAAAAAGATNAATGACAGATAATTA 120
Db 92 TTTTAAGAAATAACATTTAGTATTAGAAAAAACTAGAAAAAAGATAAATGACAGATAATTA 151
QY 121 ACTTACATGAAAAAGAAAAATATATACAAGAGACTGACAACGTATTAATTTGAATGAGA 180
Db 152 ACTTACATGAAAAAGAAAAATATATACAAGAGACTGACAACGTATTAATTTGAATGAGA 211
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAACCTTATTGTTCAATTATTCTTAATGATGG 240
Db 212 TTATAATTGAAAACTGCATCTGAAAGCAAACCTTATTGTTCAATTATTCTTAATGATGG 271
QY 241 TGTTTTATGACTAATACTGATGATTTTTCATTAAGAAACCAGTGTAAAAATATTTTAT 300
Db 272 TGTTTTATGACTAATACTGATGATTTTTCATTAAGAAACCAGTGTAAAAATATTTTAT 331
QY 301 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTCTTTTATTTGATTTGGAAGA 360
Db 332 TTTAAAAATAAGCCTGTGTGTCAAGCTCTGATCATATTCTTTTATTTGATTTGGAAGA 391
QY 361 AAATACCTGTTCTGATGACATGAATGCAAAATTTTGAATTTTAACTCNCATAATTT 420
Db 392 AAATACCTGTTCTGATGACATGAATGCAAAATTTTGAATTTTAACTCNCATAATTT 451
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCACACACTAATTAATCTGCCAGCT 480
Db 452 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCACACACTAATTAATCTGCCAGCT 511
QY 481 GTTGCAATTGTGTTTCTTACTTACTTGTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 540
Db 512 GTTGCAATTGTGTTTCTTACTTACTTGTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 571
QY 541 NGATATA-CCTTAAATATACTTTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT 598
Db 572 AGAATAATCCTTAAATATACTTTGTTAGCAAAACAAAAGCTTTTGTGTTACATAGTTCT 631
QY 599 TTGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCCATTTTCCCGACACATAATTAC 658
Db 632 TTGATTTTACTGTTCTCTAATTTTATTTCTGAAACTCAATTTTACCCAGACCAATATTAC 691
QY 659 CCTATTAACTTTGTTATGACACAGTTGT 686
Db 692 CATA-TTAACTTTGTTATGACACAGTTGT 718

RESULT 10
BU627147 762 bp mRNA linear EST 23-SEP-2002
LOCUS UI-H-FG0-bct-i-12-0-UI.s1 NCI CGAP EN1 2 Homo sapiens cDNA clone
DEFINITION UI-H-FG0-bct-i-12-0-UI 3', mRNA sequence.
ACCESSION BU627147
VERSION BU627147.1 GI:23293361
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 762)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA

sequence: 1-43, >AT rich#low complexity (matched complement)
309-335, >AT rich#Low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
Source

1.762
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FG0-bcc-1-12-0-UI"
/issue_type="Enchondroma cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP EN1_2"
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP EN1_2 is a cDNA library containing the following tissue(s): Enchondroma cell line (2 cell lines). The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is CCGGTCACTC. The cell lines was provided by Dr James Martin from University of Iowa.
TAG_LIB=UI-H-FG0
TAG_TISSUE=Enchondroma cell line (mix of EN1 and EN2)
TAG_SEQ=CGGTCACTC"

BASE COUNT 269 a 109 c 97 g 286 t 1 others
ORIGIN

Query Match 92.5%; Score 635.2; DB 13; Length 762;
Best Local Similarity 97.8%; Pred. No. 1e-65;
Matches 673; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 1 ATTTTAATTTCTATTAAACATTTCTCAAAAGCATTATTATTCCTATATCTCACTGAA 60
Db 25 ATTTTAATTTCTATTAAACATTTCTCAAAAGCATTATTATTCCTATATCTCACTGAA 84
QY 61 TTTTAAANAATAACATTAGTATTAGAAAAAAGTAAAAAAGATNAATGCAGATAATTAA 120
Db 85 TTTTAAAGAAATAACATTAGTATTAGAAAAAAGTAAAAAAGATNAATGCAGATAATTAA 144
QY 121 ACTTACATGAAAAAGAAATTTATAAAGAAAGTGAAGAGCTTATAAATGGAATGAGA 180
Db 145 ACTTACATGAAAAAGAAATTTATAAAGAAAGTGAAGAGCTTATAAATGGAATGAGA 204
QY 181 TTATTAATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATTATCTTAATGATGG 240
Db 205 TTATTAATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATTATCTTAATGATGG 264
QY 241 TGTTTATGACTAATACACTGATTTTCAATAAGAAACCAGTTAAAAATATTTTAT 300
Db 265 TGTTTATGACTAATACACTGATTTTCAATAAGAAACCAGTTAAAAATATTTTAT 324
QY 301 TTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTAATTGTTGGGAAGA 360
Db 325 TTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTAATTGTTGGGAAGA 384
QY 361 AAATACTGTTTCTGATAGCAATAAGCAAAATTTTAGATTTTAACTCNCATAATTT 420
Db 385 AAATACTGTTTCTGATAGCAATAAGCAAAATTTTAGATTTTAACTCNCATAATTT 444
QY 421 AAGAACTATTGAGAATTGATTAATGACATGAAGTGACACAACTAATTACTGGCCAGCT 480
Db 445 AAGAACTATTGAGAATTGATTAATGACATGAAGTGACACAACTAATTACTGGCCAGCT 504
QY 481 GTTGGCATGTGTTCTTACTAGTTCTCCCAAGAAACCTTTAAACTGAATCTTCAGC 540
|||||

Db 505 GTTGGCATGTGTTCTTACTAGTTCTCCCAAGAAACCTTTAAACTGAATCTTCAGC 564
QY 541 NGATATA-CCTTAATATACCTTTAGCCAAACAAA-CCTTTTGTGTACATAGTTCT 598
Db 565 AGATAATCCTTAATATACCTTTAGCAAAACAAAGCTTTTGTGTACATAGTTCT 624
QY 599 TTGATTTTACTGTTCCCTAATTTTATTTCTGAAACTCCATTTTCCCAGACCATATATAC 658
Db 625 TTGATTTTACTGTTCCCTAATTTTATTTCTGAAACTCAATTTTACCAGACCAATATATAC 684
QY 659 CCTATTTAATCTTGTATGACAGTTGT 686
Db 685 CATATTTAATCTTGTATGACAGTTGT 711
|||||

RESULT 11

LOCUS A1826287 770 bp mRNA linear EST 21-DEC-1999
DEFINITION wk33fe07.x1 NCI CGAP Pr22 Homo sapiens cDNA clone IMAGE:2417221 3' similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (HUMAN); contains element MER28 repetitive element ;, mRNA sequence.

ACCESSION A1826287 GI:5446958
VERSION A1826287.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 770)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished

JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www-bio.lnl.gov/bbrp/image/image.html
Insert length: 1454 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 404.

FEATURES
Source

1.770
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2417221"
/sex="male"
/issue_type="normal prostate"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr22"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 271 a 111 c 99 g 285 t 4 others
ORIGIN

Query Match 92.3%; Score 634.4; DB 9; Length 770;
Best Local Similarity 97.5%; Pred. No. 1.3e-65;
Matches 671; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 1 ATTTTAATTTCTATTAAACATTTCTCAAAAGCATTATTATTCCTATATCTCACTGAA 60
|||||

Db 38 ATTTTAATTTCTATTAACATCTCTCAAGCATTATTATTACCTATATCTCACTGAA 97
QY 61 TTTTANNAATAACATAGTATAGAAAACTAGAAAAAGATNAATGCAGATAATTAA 120
Db 98 TTTTAAGAAATAACATTAGTATAGAAAACTAGAAAAAGATNAATGCAGATAATTAA 157
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGGAGCTGAGAACGTTATAAATTGAAATGAGA 180
Db 158 ACTTACATGAAAAAGAAAAATTATACAAAGGAGCTGAGAACGTTATAAATTGAAATGAGA 217
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAACTTTATGTTCAATTATCTTAATGATGG 240
Db 218 TTATAATTGAAAACTGCATCTGAAAGCAACTTTATGTTCAATTATCTTAATGATGG 277
QY 241 TGTTTATGACTAATACACTGATTTTCAATAAGGAAACCCAGTTAAAAATATTTTAT 300
Db 278 TGTTTATGACTAATACACTGATTTTCAATAAGGAAACCCAGTTAAAAATATTTTAT 337
QY 301 TTTAAAAATAAGCCTGTGTTCAAGCTGATCATATTTCTTTTAAATTGATTGGAAGA 360
Db 338 TTTAAAAATAAGCCTGTGTTCAAGCTGATCATATTTCTTTTAAATTGATTGGAAGA 397
QY 361 AAATACGTCTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATCTCNCATATTT 420
Db 398 AAATACGTCTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATCTCNCATATTT 457
QY 421 AAGACTAATTGAGAAATTGATTAAATGACATGAGTGCACAACACTAATTAATCTGGCCAGCT 480
Db 458 AAGACTAATTGAGAAATTGATTAAATGACATGAGTGCACAACACTAATTAATCTGGCCAGCT 517
QY 481 GTTGGCATGTGTTCTTACTAGTCTCCCAAGGAAACCTTTAACTGAATCTTCAGC 540
Db 518 GTTGGCATGTGTTCTTACTAGTCTCCCAAGGAAACCTTTAACTGAATCTTCAGC 577
QY 541 NGAATAA-CCTTAATATATCTTTGTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT 598
Db 578 AGAATATCTCTTAATATATCTTTGTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT 637
QY 599 TTGGATTTTACTGTTCTCTAATTTTATTTGAAACTCCATTTTCCCGACAGCATATATAC 658
Db 638 TTGGATTTTACTGTTCTCTAATTTTATTTGAAACTCAATTTTACCAGACCATATATAC 697
QY 659 CCTATTAACTTTGTTATGCACAGTTGT 686
Db 698 CATA-TTAACTNTGTAATGCACAGTTGT 724

RESULT 12
LOCUS BM968746 721 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-CF-DUI-aam-a-03-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone
ACCESSION BM968746
VERSION BM968746.1 GI:19586333
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 721)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 315-341, >AT-rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source

location/Qualifiers
1..721
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-aam-a-03-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

BASE COUNT
ORIGIN

257 a 97 c 91 g 276 t

Query Match 92.0%; Score 632; DB 12; Length 721;
Best Local Similarity 97.5%; Pred. No. 2.5e-65;
Matches 671; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 1 ATTTTAATTTCTATTAACATCTCTCAAGCATTATTATTACCTATATCTCACTGAA 60
Db 31 ATTTTAATTTCTATTAACATCTCTCAAGCATTATTATTACCTATATCTCACTGAA 90
QY 61 TTTTANNAATAACATAGTATAGAAAACTAGAAAAAGATNAATGCAGATAATTAA 120
Db 91 TTTTAAGAAATAACATTAGTATAGAAAACTAGAAAAAGATNAATGCAGATAATTAA 150
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGGAGCTGAGAACGTTATAAATTGAAATGAGA 180
Db 151 ACTTACATGAAAAAGAAAAATTATACAAAGGAGCTGAGAACGTTATAAATTGAAATGAGA 210
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAACTTTATGTTCAATTATCTTAATGATGG 240
Db 211 TTATAATTGAAAACTGCATCTGAAAGCAACTTTATGTTCAATTATCTTAATGATGG 270
QY 241 TGTTTATGACTAATACACTGATTTTCAATAAGGAAACCCAGTTAAAAATATTTTAT 300
Db 271 TGTTTATGACTAATACACTGATTTTCAATAAGGAAACCCAGTTAAAAATATTTTAT 330
QY 301 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTAAATTGATTGGAAGA 360
Db 331 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTAAATTGATTGGAAGA 390
QY 361 AAATACGTCTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATCTCNCATATTT 420
Db 391 AAATACGTCTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATCTCNCATATTT 450

QY 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGACACAACTAATTACTGGCCAGCT 480
|||||
Db 451 AAGAACTATTGAGAAATGATTAATGACATGAGTGACACAACTAATTACTGGCCAGCT 510
481 GTGGCATTGTGTTTCTTACTAGTCTCCCAAGAAAACCTTAACTGAATCTTCAGC 540
|||||
Db 511 GTTGGCATGTGTTCTTACTAGTCTCCCAAGAAAACCTTAACTGAATCTTCAGC 570
541 NGAAATAA-CCTTAATATACTTGTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT 598
|||||
Db 571 AGAATAATCCTTAATATACTTGTGTTAGCCAAACAAAACCTTTTGTGTTACATAGTTCT 630
QY 599 TTGGATTTTACTGTTCCCTAATTATTTCTGAAACTCCATTTTCCCGACAGACCATTAATTAC 658
|||||
Db 631 TTGGATTTTACTGTTCCCTAATTATTTCTGAAACTCAATTTTACCCGACAGACCATTAATTAA 690
QY 659 CCTATTTAACCTTGTATGACAGTTGT 686
|||
Db 691 CATATA-TTAACTTTGGTATGACAGTTGT 717

RESULT 13
CA424156 734 bp mRNA linear EST 07-NOV-2002
LOCUS
DEFINITION UI-H-FEI-bdv-m-12-0-UI.s1 NCI_CGAP_FEI Homo sapiens cDNA clone
UI-H-FEI-bdv-m-12-0-UI 3', mRNA sequence.

ACCESSION CA424156
VERSION CA424156.1 GI:24786882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-romail.nih.gov
Tissue Procurement: James Martin
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-48, >AT-rich#low complexity (matched complement)
314-340, >AT-rich#low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
Source
Location/Qualifiers
1..734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-bdv-m-12-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FEI"
/note="Organ: Chondrosarcoma; Vector: p77T3-Pac (Pharmacia
) with a modified polylinker; Site 1: Ecor I; Site 2: Not
I; NCI_CGAP FEI is a normalized cDNA library derived from
a pool of mRNA obtained from 3 cell lines from grade II
chondrosarcoma tissues. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into p77T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGCTACGGAC. The cell
lines were provided by Dr James Martin from the University
of Iowa.
TAG_LIB=UI-H-FEI
TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG_SEQ=CGCTACGGAC"
BASE COUNT 261 a 101 c 94 g 277 t 1 others
ORIGIN

Query Match 92.0%; Score 632; DB 14; Length 734;
Best Local Similarity 97.5%; Pred. No. 2.5e-65;
Matches 671; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 1 ATTTAATTTCTATTAACAATCTCTCAAGCATTAATTATCTATATCTACCTGAA 60
|||||
Db 30 ATTTAATTTCTATTAACAATCTCTCTCAAGCATTAATTATCTATATCTACCTGAA 89
61 TTTAANAATAACATTAATGATTAAGAAAAGTAAAGTAAATGACATTAATTAA 120
|||||
Db 90 TTTAAGAATAACATTAATGATTAAGAAAAGTAAAGTAAATGACATTAATTAA 149
QY 121 ACTTACATGAAGAAAGAAATTTATACAAAGAGCTGAGAACCTTATAATTGAATGAGA 180
|||||
Db 150 ACTTACATGAAGAAAGAAATTTATACAAAGAGCTGAGAACCTTATAATTGAATGAGA 209
181 TTATAATTGAAAACCTGCATCTGAAAAGCAAACTTTATGTTCAATTATTCTTAATGATGG 240
|||||
Db 210 TTATAATTGAAAACCTGCATCTGAAAAGCAAACTTTATGTTCAATTATTCTTAATGATGG 269
QY 241 TGTTTTATGACTAATATACATGATTTTTCATTAAGAAAACCATGTTAAATAATTTTAT 300
Db 270 TGTTTTATGACTAATATACATGATTTTTCATTAAGAAAACCATGTTAAATAATTTTAT 329
QY 301 TTTAAAAATAGCCCTGTGTTCAAGCTCTGATCATATTCTTTATTTTGGATTTGGAGA 360
|||||
Db 330 TTTAAAAATAGCCCTGTGTTCAAGCTCTGATCATATTCTTTATTTGGATTTGGAGA 389
QY 361 AATACTGTTCTGATAGCATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420
|||||
Db 390 AATACTGTTCTGATAGCATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 449
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGACACAACTAATTACTGGCCAGCT 480
450 AAGAACTATTGAGAAATGATTAATGACATGAGTGACACAACTAATTACTGGCCAGCT 509
481 GTGGCATTGTGTTTCTTACTAGTCTCCCAAGAAAACCTTAACTGAATCTTCAGC 540
|||||
Db 510 GTTGGCATGTGTTTCTTACTAGTCTCCCAAGAAAACCTTAACTGAATCTTCAGC 569
541 NGAAATAA-CCTTAATATACTTGTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT 598
|||||
Db 570 AGAATAATCCTTAATATACTTGTGTTAGCCAAACAAAACCTTTTGTGTTACATAGTTCT 629
QY 599 TTGGATTTTACTGTTCCCTAATTATTTCTGAAACTCCATTTTCCCGACAGACCATTAATTAC 658
|||||
Db 630 TTGGATTTTACTGTTCCCTAATTATTTATTTCTGAAACTCAATTTTACCCCGACAGACATTAATTAC 689
QY 659 CCTATTTAACCTTGTATGACAGTTGT 686
|||
Db 690 CATATA-TTAACTTTGGTATGACAGTTGT 716

RESULT 14
AW052045 786 bp mRNA linear EST 20-SEP-1999
LOCUS
DEFINITION wx25e05.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2544704 3',
similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION AW052045
VERSION AW052045.1 GI:5914404
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 458.
location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2544704"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
plasmid DNA from the normalized library NCI_CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 290 a 116 c 109 g 270 t 1 others
ORIGIN
Query Match 92.0%; Score 632; DB 9; Length 786;
Best Local Similarity 97.5%; Pred. No. 2.4e-65;
Matches 671; Conservative 0; Mismatches 14; Indels 3; Gaps 3;
OY 1 ATTTAAATTTCTATTAACATCTCTCAAGCATTATTTTACCTATATCTCACTGAA 60
|||||
Db 15 ATTTAAATTTCTATTAACATCTCTCGGAAAGCATTATTTTACCTATATCTCACTGAA 74
OY 61 TTTTAAATAATACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 120
|||||
Db 75 TTTTAAATAATACATTAGTATTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 134
OY 121 ACTTACATGAAAAAGAAAAATTATACAAAGGACTGAGAACGTTATAATTGAAATGAGA 180
|||||
Db 135 ACTTACATGAAAAAGAAAAATTATACAAAGGACTGAGAACGTTATAATTGAAATGAGA 194
OY 181 TTATAATTGAAAACTGCATCTGAAAAGCAACTTATTGTTCAATTATTCTTAATGATGG 240
|||||
Db 195 TTATAATTGAAAACTGCATCTGAAAAGCAACTTATTGTTCAATTATTCTTAATGATGG 254
OY 241 TGTTTTATGACTAATACACTGATTTTTCATAAGAAACCCATGTTAAAAATATTTTAT 300
|||||
Db 255 TGTTTTATGACTAATACACTGATTTTTCATAAGAAACCCATGTTAAAAATATTTTAT 314
OY 301 TTTAAATAATAGCCTGTGTCAAGCTCTGATCATATTCTTTAATTGATTGGAGAGA 360
|||||
Db 315 TTTAAATAATAGCCTGTGTCAAGCTCTGATCATATTCTTTAATTGATTGGAGAGA 374
OY 361 AAATACGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATCTCNCTAATTTT 420
|||||
Db 375 AAATACGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTTAATCTCACTAATTTT 434

OY 421 AAGACTATTGAGAAATGATTAATGACATGAAGTGACACACACTAATTACTGGCCAGCT 480
|||||
Db 435 AAGACTATTGAGAAATGATTAATGACATGAAGTGACACACACTAATTACTGGCCAGCT 494
OY 481 GTTGGCATGTGTTTCTTACTTAGTCTCCCAAGGAAACTCTTAACTGAATCTTCAGC 540
|||||
Db 495 GTTGGCATGTGTTTCTTACTTAGTCTCCCAAGGAAACTCTTAACTGAATCTTCAGC 554
OY 541 NGATAAA-CCTTAATATACTTTGTTAGCCAAACAAA-CTTTTGTGTACATAGTTCT 598
|||||
Db 555 AGAATATATCCTTAATATACTTTGTTAGCCAAACAAAAGCTTTTGTGTACATAGTTCT 614
OY 599 TTGATTTTACTGTTCTTAATTTTATTCTGAAACTCCATTTTCCCGAGACCATATTAC 658
|||||
Db 615 TTGATTTTACTGTTCTTAATTTATTTCTGAAACTCAATTTTACCAGACCATATTAC 674
OY 659 CCTAATTAACTTTGTTATGACAGTTGT 686
|||||
Db 675 CATA-TTAACTTTGTATGACAGTTGT 701

RESULT 15
BQ009270 750 bp mRNA linear EST 26-MAR-2002
LOCUS
DEFINITION UT-H-ED1-axx-n-02-0-UT.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone
IMAGE:5835073 3', mRNA sequence.
ACCESSION BQ009270.1 GI:19734171
VERSION BQ009270
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 750)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
The following repetitive elements were found in this cDNA
sequence: 316-342, >AT-rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5835073"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/note="Organ: left Pubic Bone; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site 1: EcoR I;
Site 2: Not I; NCI_CGAP_ED1 is a normalized cDNA library
containing the following tissue(s): Chondrosarcoma cell
line C55. The library was constructed according to Bonaldo
, Lennon and Soares, Genome Research, 6:791-806, 1996.
First strand cDNA synthesis was primed with an oligo-dT
primer containing a Not I site. Double stranded cDNA was
ligated to an EcoR I adaptor, digested with Not I, and
cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The

FEATURES
source

sequence tag for this library is GCTCAAGGCT.

TAG_LIB=UI-H-ED1

TAG_TISSUE=chondrosarcoma

TAG_SEQ=CGTCAAGGCT"

BASE COUNT 266 a 105 c 96 g 281 t 2 others

ORIGIN

```
Query Match          91.7%; Score 630.2; DB 12; Length 750;
Best Local Similarity 96.4%; Pred. No. 4e-65;
Matches 663; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

QY      1 ATTTTAATTTCTATTAACATCTCTCTCAAGCATTTATCTATATCTCACTGAA 60
      |||
Db      32 ATTTTAATTTCTATTAACATCTCTCTCAAGCATTTATCTATATCTCACTGAA 91

QY      61 TTTTAAATAATACATAGTATTAGAAAACTAGGAAAAAGATNAATGCAGATAATTAA 120
      |||
Db      92 TTTTAAGAAATTAACATTAGTATTAGAAAACTAGGAAAAAGATNAATGCAGATAATTAA 151

QY      121 ACTTACATGAAAAAGAAAAATTATACAAAGGACTGAGAACGTTATAATTGAAATGAGA 180
      |||
Db      152 ACTTACATGAAAAAGAAAAATTATACAAAGGACTGAGAACGTTATAATTGAAATGAGA 211

QY      181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 240
      |||
Db      212 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 271

QY      241 TGTTTATGACTAATACACTGATTTTTCATAGGAAACCACTGTTAAATATTTTAT 300
      |||
Db      272 TGTTTATGACTAATACACTGATTTTTCATAGGAAACCACTGTTAAATATTTTAT 331

QY      301 TTTAAAAATAAGCCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATTGGGAAGA 360
      |||
Db      332 TTTAAAAATAAGCCCTGTGTCAAGCTCTGATCATATTTCTTTATTTTGATTGGGAAGA 391

QY      361 AAATACTGTTTCTGATAGCATGAAATGCAAAATTTTAGATTTTAACTCNCNTAATTTT 420
      |||
Db      392 AAATACTGTTTCTGATAGCATGAAATGCAAAATTTTAGATTTTAACTCNCNTAATTTT 451

QY      421 AAGAACTATTGAGAAATGATTAATGACATGAAAGTGACACAACACTAATTACTGGCCAGCT 480
      |||
Db      452 AAGAACTATTGAGAAATGATTAATGACATGAAAGTGACACAACACTAATTACTGGCCAGCT 511

QY      481 GTTGGCATGTGTTTCTTACTTAGTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 540
      |||
Db      512 GTTGGCATGTGTTTCTTACTTAGTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 571

QY      541 NGAATAA-CCTTAATATACTTTGTTAGCCCAACAAAAA-CCTTTTGTGTTTACATAGTTCT 598
      |||
Db      572 AGAATAAATCCTTAATATACTTTGTTAGCAAAACAAAAAGCTTTTGTGTTTACATAGTTCT 631

QY      599 TTGATTTTACTGTTCTCTAATTTTATCTGAAACTCCATTTTCCCGAGACCATTAATTAC 658
      |||
Db      632 TTGATTTTACTGTTCTCTAATTTTANCTGAAACTCAATTTTACCCCAAGACCATATATTA 691

QY      659 CCTATTTAACCTTGTATGACACAGTTGT 686
      |||
Db      692 CCATATTTAACCTTGTATGACACAGTTGT 719
```

Search completed: November 27, 2003, 12:28:12
Job time : 1852.59 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:01:24 ; Search time 225.747 Seconds
(without alignments)
8215.019 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 687

Sequence: 1 atttcaatttctataaac.....cttgctatgcacagttgtc 687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	683	99.4	687	20	AAX40095	Gastric cancer aas
c 2	636.8	92.7	2663	22	AAH57501	Human liver cell s
3	611.8	89.1	714	20	AAX40094	Gastric cancer aas
4	467.8	68.1	486	24	ABV87267	Human colon cancer
5	411.8	59.9	843	20	AAX40097	Gastric cancer aas
c 6	300.8	43.8	351	16	AAT19975	Human gene signatu
c 7	261	38.0	262	22	AAH57293	Human liver specif
c 8	161	23.4	173	25	ABX83093	Corn ear-derived p

c 9	126.2	18.4	263	15	AAQ76489	Human genome fragm
10	103.6	15.1	397	25	ABX40797	Bovine EST associa
c 11	85.8	12.5	2152	24	AB199762	Mouse ischaemic co
12	71.6	10.4	8056	25	ABZ10246	Haematopoietic cel
c 13	70	10.2	8056	25	ABZ10246	Haematopoietic cel
14	66.8	9.7	7676	24	ABL70409	Chemically treated
15	66.8	9.7	7676	24	ABL34598	Human metastasis a
16	66.8	9.7	8056	25	ABZ10100	Haematopoietic cel
c 17	66.4	9.7	5852	12	AAQ11710	Diclyostelium plas
18	65.4	9.5	10048	24	ABL70313	Chemically treated
c 19	65.4	9.5	10048	24	AAS61251	Human gene regulat
c 20	65.2	9.5	8056	25	ABZ10100	Haematopoietic cel
c 21	65	9.5	113515	24	ABL34174	Human immune syste
c 22	63.4	9.2	883	22	ABL15210	Human breast cance
c 23	63	9.2	6106	22	AAS46429	Tumour suppressor
c 24	63	9.2	6106	24	ABK40031	Human chemically p
c 25	63	9.2	6106	24	ABL33472	Human immune syste
c 26	63	9.2	13131	24	ABL92249	Chemically treated
c 27	61.8	9.0	291	25	ABX48529	Bovine EST associa
c 28	60.8	8.9	6699	24	AAS63321	Chemically pretrea
c 29	60.4	8.8	6169	22	AAS46370	Tumour suppressor
c 30	60.4	8.8	6169	24	ABN80097	Human DNA for stag
c 31	60.4	8.8	6944	24	ABK34026	Human chemically m
c 32	60.2	8.8	6509	24	ABL32226	Human immune syste
c 33	60.2	8.8	6912	24	ABK28371	DNA transcription
c 34	59.8	8.7	7459	24	ABK31383	Signal transductio
c 35	59.8	8.7	17721	24	ABL33729	Human immune syste
c 36	59.6	8.7	6175	24	AAD28372	Human chemically t
c 37	59.6	8.7	10467	24	ABK28453	DNA transcription
c 38	59.4	8.6	3991	22	AAD16633	Human novel protei
c 39	59.2	8.6	5815	24	ABK40024	Human chemically p
c 40	59.2	8.6	640681	24	ABA92787	Buchnera sp. genom
c 41	58.8	8.6	449	25	ABX48095	Bovine EST associa
c 42	58.8	8.6	6071	24	ABL92215	Chemically treated
c 43	58.8	8.6	6071	24	AAD23316	Chemically treated
c 44	58.4	8.5	424	25	ABX46053	Bovine EST associa
c 45	58.2	8.5	5962	24	ABL33287	Human immune syste

ALIGNMENTS

RESULT 1						
ID	AAX40095	standard;	DNA;	687	BP.	
AC	AAX40095;					
XX						
DT	02-JUL-1999	(first entry)				
XX						
DE	Gastric cancer associated gene.					
XX						
KW	Cancer associated antigen; diagnosis; research; treatment; human;					
KW	breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;					
KW	prostate cancer; ss.					
XX						
OS	Homo sapiens.					
XX						
PN	WO9904265-A2.					
XX						
PD	28-JAN-1999.					
XX						
PF	15-JUL-1998;	98NO-US14679.				
XX						
PR	22-JUN-1998;	98US-0102322.				
PR	17-JUL-1997;	97US-0896164.				
PR	10-OCT-1997;	97US-0061599.				
PR	10-OCT-1997;	97US-0061765.				
PR	10-OCT-1997;	97US-0948705.				
PR	11-OCT-1997;	97GB-0021697.				
XX						
PA	(LUDW-) LUDWIG INST CANCER RES.					
XX						

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ,
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E,
PI Tureci O;

WPI; 1999-132448/11.

PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
PS Claim 67; Page 695; 787bp; English.

Claim 67; Page 695; 787pp; English.

The invention relates to a method for diagnosing a disorder characterised by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule (NAM). The method comprises: (a) contacting a biological sample isolated from a subject with an agent that specifically binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the interaction between the agent and the NAM or the expression product as a determination of the disorder. The products and methods can be used in the diagnosis, monitoring, research, or treatment of conditions characterised by the expression of various cancer associated antigens. The invention provides nucleic acid sequences and encoded polypeptides which are cancer associated antigen precursors expressed in human breast cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and lung cancer.

SQ Sequence 687 BP; 242 A; 98 C; 85 G; 258 T; 4 other;

Query Match	99.4%;	Score 683;	DB 20;	Length 687;
Best Local Similarity	100.0%;	Pred. No. 4.6e-108;		
Matches 687; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 1 ATTTTAATTTCTATTAAACATTCTTCTCAAGCATTATTTATCCATAATCTCACTGAA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 ATTTTAATTTCTATTAAACATTCTTCTCAAGCATTATTTATCCATAATCTCACTGAA 60

[illegible][illegible]

Qy	181 TTATAATTTGAAAACTGCATCTGAAGCAAATTATTTGTTC AATTATTC TTAATGATGG 240
Db	181 TTATAATTTTGA AAACTGCATCTGAAGCAAATTATTTGTTC AATTATTC TTAATGATGG 240

[illegible]

Oy 301 TTTAAAAATAGCCCTGTGTTCAAGCCTCGATCATATTTCCTTATTGATTGGGAAGA 360

Db 301 TTTAAAAATAGCCCTGTGTTCAAGCCTCGATCATATTTCCTTATTGATTGGGAAGA 360

```
Oy      361 AAATACGTCTTCTGATAGCATGAATGC AAAATT TTTAGATTTT TAATCTCNC TAAATTTT   420
         |||||
Db       361 AAATACTGTTTTCTGATAGCATGA AATGC AAAAATTT TT TAGATTTT TAATCTC NC TAAATTTT   420
```

Oy	421 AAGAACTATTGAGAAATTCATTAAATGCATGAAGTGCACAACACTAATTACTGGCCAGCT	480
Db	421 AAGAACTATTGAGAAATTCATTAAATGCATGAAGTGCACAACACTAATTACTGGCCAGCT	480

[illegible]

Qy	541	NGAATAACCTTAATATACCTTTGTTAGCCAAACAAA	CTTTTGTTCATAGTTCCTT	6000
Db	541	NGAATAACCTTAATATACCTTTGTTAGCCAAACAAA	CTTTTGTTCATAGTTCCTT	6000

[illegible]

Qy	661 TATTAACTTGTGTTANGACAGTTGTT	6877
Db	661 TATTAACTTGTGTTATGCACAGTTGTT	6877

RESULT 2
AAH57501/C

AC AAH57501;

DT 10-SEP-2001 (first entry)

DE Human liver cell specific cDNA sequence SEQ ID NO:341.

KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
 KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

Os Homo sapiens.

PN WO200132927-A2.

PD 10-MAY-2001.

PF 02-NOV-2000; 2000WO-US30396.

PR 04-NOV-1999; 99US-0163508.

PA (INCY-) INCYTE GENOMICS INC.

PI Sornasse T, Seilhamer JJ, Watson GA;

DR WPI; 2001-291057/30.

PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -

PS Claim 1; Page 258-259; 327pp; English.

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide sequences (I). (I) can have cytostatic, immunomodulatory and neuroprotective activities, and can be used in gene therapy. (I) and proteins (II) encoded by then are used in high throughput screening assays to select DNA molecules, RNA molecules, peptide nucleic acids, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents. Expression of (I) in a sample indicates the differentiation of embryonic stem cells into a tissue selected from brain, heart, kidney, liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used to produce an expression profile that defines a metabolic or developmental process, treatment, condition, disease or disorder. The gene profile can be used for diagnosis, prognosis or monitoring of treatments and for investigating a predisposition to a disorder where the gene is associated with a cancer, immunopathology or neuropathology.

SQ Sequence 2663 BP; 807 A; 472 C; 599 G; 785 T; 0 other;

Query Match	92.7%;	Score 636.8;	DB 22;	Length 2663;
Best Local Similarity	98.0%;	Pred. No. 3.4e-100;		
Matches 674;	Conservative 0;	Mismatches 11;	Indels 3;	Gaps 3;

Qy	1	ATTTTAATTTCTATTAACATTCTTCTCAAGCATTATTTATCCATAATCTCACTGAA	60
Db	2639	ATTTTAATTTCTATTAACATTCTTCTCAAGCATTATTTATCCATAATCTCACTGAA	2580

QY	61	TTTTAANAATAACATTAGTATTAGAAAAAAGATNAATGACATAATTAA	120
DB	2579	TTTTAAGAATAACATTAGTATTAGAAAAAAGATNAATGACATAATTAA	2520
QY	121	ACTTACATGAAAAAGAAAAATTATACAAAGACGTGAGACGTTATAAATGAAATGAGA	180
DB	2519	ACTTACATGAAAAAGAAAAATTATACAAAGACGTGAGACGTTATAAATGAAATGAGA	2460
QY	181	TTATATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATCTTAATGATGG	240
DB	2459	TTATATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATTATCTTAATGATGG	2400
QY	241	TGTTTATGACTTAATACACTGATTTTCAATAGAAACCCATGTTAAAAATATTTTAT	300
DB	2399	TGTTTATGACTTAATACACTGATTTTCAATAGAAACCCATGTTAAAAATATTTTAT	2340
QY	301	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTTTGGGGAAGA	360
DB	2339	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTTGGAAGA	2280
QY	361	AAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCATAATTT	420
DB	2279	AAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCATAATTT	2220
QY	421	AAGAACTATTGAGAAATTGATTAATGACATGAGGCAACAACATAATCTGCGCAGCT	480
DB	2219	AAGAACTATTGAGAAATTGATTAATGACATGAGGCAACAACATAATCTGCGCAGCT	2160
QY	481	GTTGGCATTGTGTTCTTACTTAGTCTCCCAAGAGAAACTCTTAACTGAACTCTCAGC	540
DB	2159	GTTGGCATTGTGTTCTTACTTAGTCTCCCAAGAGAAACTCTTAACTGAACTCTCAGC	2100
QY	541	NGAATAA-CCCTTAATATATACCTTTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT	598
DB	2099	AGAATAATCCTTAATATATACCTTTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTCT	2040
QY	599	TTGGATTTTACTGTTCTTAATTTTATCTGAACCTCCATTTTCCCGACAGCATTAATTAC	658
DB	2039	TTGGATTTTACTGTTCTTAATTTTATCTGAACCTCCATTTTCCCGACAGCATTAATTAC	1980
QY	659	CCTATTTAACCTTGTATTGACACAGTTGT	686
DB	1979	CATA-TTAACTTTGTTAATGCACAGTTGT	1953
RESULT 3			
AAK40094			
ID	AAK40094 standard; DNA; 714 BP.		
XX			
AC	AAK40094;		
XX			
DT	02-JUL-1999 (first entry)		
XX			
DE	Gastric cancer associated gene.		
XX			
KW	Cancer associated antigen; diagnosis; research; treatment; human;		
KW	breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;		
KW	prostate cancer; ss.		
OS	Homo sapiens.		
XX			
PN	W09904265-A2.		
XX			
PD	28-JAN-1999.		
XX			
PF	15-JUL-1998; 98WO-US14679.		
XX			
PR	22-JUN-1998; 98US-0102322.		
PR	17-JUL-1997; 97US-0896164.		
PR	10-OCT-1997; 97US-0061599.		
PR	10-OCT-1997; 97US-0061765.		
PR	10-OCT-1997; 97US-0948705.		
PR	11-OCT-1997; 97GB-0021697.		

XX	(LUDW-) LUDWIG INST CANCER RES.	
PA		
XX		
PI	Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;	
PI	Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;	
PI	Tureci O;	
XX		
DR	WPI; 1999-132448/11.	
XX		
PT	New isolated cancer associated nucleic acids and polypeptides -	
PT	isolated using sera from cancer patients, used to develop products	
PT	for the diagnosis, monitoring or treatment of cancers	
XX		
PS	Claim 67; Page 695; 787pp; English.	
XX		
CC	The invention relates to a method for diagnosing a disorder characterised	
CC	by expression of a human cancer associated antigen precursor coded for by	
CC	a nucleic acid molecule (NAM). The method comprises: (a) contacting a	
CC	biological sample isolated from a subject with an agent that specifically	
CC	binds to the NAM, an expression product or a fragment of an expression	
CC	product complexed with an HLA molecule; and (b) determining the	
CC	interaction between the agent and the NAM or the expression product as a	
CC	determination of the disorder. The products and methods can be used in	
CC	the diagnosis, monitoring, research, or treatment of conditions	
CC	characterised by the expression of various cancer associated antigens.	
CC	The invention provides nucleic acid sequences and encoded polypeptides	
CC	which are cancer associated antigen precursors expressed in human breast	
CC	cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and	
CC	lung cancer.	
XX		
SQ	Sequence 714 BP; 258 A; 100 C; 92 G; 260 T; 4 other;	
	Query Match	89.1%; Score 611.8; DB 20; Length 714;
	Best Local Similarity	97.1%; Pred. No. 6.5e-96;
	Matches 671; Conservative	0; Mismatches 15; Indels 5; Gaps 5;
QY	1	ATTTAATTTCTATTAACATTTCTCAAGCATTTATTTATCCATATCTCACTGAA 60
DB	15	ATTTAATTTCTATTAACATTTCTCAAGCATTTATTTATCCATATCTCACTGAA 74
QY	61	TTTTAANAATAACATTAGTATTGAGAAAACTAGGAAAAAGATNAATGACATAATTAA 120
DB	75	TTTTAAGAAATAACATTAGTATTGAGAAAACTAGGAAAAAGATNAATGACATAATTAA 134
QY	121	ACTTACATGAAAAAGAAAAATTATACAAAGACTGAGAACGTTATTAATTGAAATGAGA 180
DB	135	ACTTACATGAAAAAGAAAAATTATACAAAGACTGAGAACGTTATTAATTGAAATGAGA 194
QY	181	TTATAATTTGAAAACTGCATCTGAAGCAAACTTTATTGTTCAATTAATCTTAATGATGG 240
DB	195	TTATAATTTGAAAACTGCATCTGAAGCAAACTTTATTGTTCAATTAATCTTAATGATGG 254
QY	241	TGTTTATGACTAATACACTGATTTTTCATAAGGAAACCCATGTTAAAAATATTTTAT 300
DB	255	TGTTTATGACTAATACACTGATTTTTCATAAGGAAACCCATGTTAAAAATATTTTAT 314
QY	301	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTTGATTGGGAAGA 360
DB	315	TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTTGATTGGGAANA 374
QY	361	AAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCATAATTTT 420
DB	375	AAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCATAATTTT 434
QY	421	AAGAACTATTGAGAAATTGATTAATGACATGAAGTGACAAACATAATTAATGAGCCAGCT 480
DB	435	AANAACATTTGAGAAATTGATTAATGACATGAAGTGACAAACATAATTAATGAGCCAGCT 494
QY	481	GTTGGCATTTGTGTTCTTACTTAGTTCTCCCAA-GGAAAACTCTTAACCTGAATCTTCAG 539
DB	495	GTTGGCATTTGTGTTCTTACTTAGTTCTCCCAAAGGAAAACTCTTAATGGAATCTTCAG 554
QY	540	CNGAATAA-CCTTAATATACTTTGTTAGCCAAACAAA-CTTTTGTGTTACATAGTTC 597

Db 555 CAGAATAATCCTTAATATACTTTGTAAGCAAAAGCTTTTGTGTTACATAGTTC 614
QY 598 TTT-GGATTCTAGTCTTCTAATTTTATTCGAACTCCATTTTCCCGACAGACCAATATT 656
Db 615 TTGGGATTCTTACTGTTCTTAATTTTATTCTGAACCTCAATTTTACCACAGACCAATATT 674
QY 657 ACCCTATTTAACCTTGTGTAAGCAGAGTTGTT 687
Db 675 ACCATA-TTAACTTGTGTTGACAGAGTTGTT 704
RESULT 4
ABV87267
ID ABV87267 standard; cDNA; 486 BP.
XX
AC ABV87267;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human colon cancer related cDNA SEQ ID NO 578.
XX
KM Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200258534-A2.
XX
PD 01-AUG-2002.
XX
PF 19-NOV-2001; 2001WO-US43704.
XX
PR 20-NOV-2000; 2000US-252222P.
PR 06-FEB-2001; 2001US-267011P.
PR 28-MAR-2001; 2001US-279670P.
PR 10-JUL-2001; 2001US-304037P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
XX
DR WPI; 2002-608400/65.
XX
PT New isolated tumor colon polynucleotide and polypeptide, useful for the
PT diagnosis, prevention and/or treatment of cancer, in particular colon
PT cancer -
XX
PS Claim 1; SEQ ID NO 578; 266pp + Sequence Listing; English.
XX
CC The invention relates to a human colon tumour expressed polynucleotide
CC (I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of
CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
CC complements of (i); (iii) at least 20 contiguous residues of (i); (iv)
CC sequences that hybridize to (i), under moderately stringent conditions;
CC (v) sequences having at least 75% or 90% identity to (i); or (vi)
CC degenerate variants of (i). The compositions, prevention and methods of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC cancer, particularly colon cancer. (I) can be used in gene therapy and
CC (II) and (III) are useful in pharmaceutical compositions such as vaccines.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 486 BP; 192 A; 55 C; 64 G; 175 T; 0 other;

Query Match 68.1%; Score 467.8; DB 24; Length 486;
Best Local Similarity 98.9%; Pred. No. 2.4e-71;
Matches 469; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATTTTAATTTCTAATAACATTCTCTCAAGCATTATTTATCTATATCTCACTGAA 60
Db 13 ATTTTAATTTCTAATAATATCTCTCTCAAGCATTATTTATCTATATCTCACTGAA 72

QY 61 TTTTAAATAATTAACATTAGTATTAGAAAACTAGAAAAAGATNATGACATTAATTAA 120
Db 73 TTTAAGAAATAACATTAGTATTAGAAAACTAGAAAAAGATNATGACATTAATTAA 132
QY 121 ACTTACATGAAAAAGAAAAATTATTAACAAGGACTGAAACGTTATTAATTGAAATGAGA 180
Db 133 ACTTACATGAAAAAGAAAAATTATTAACAAGGACTGAAACGTTATTAATTGAAATGAGA 192
QY 181 TTATTAATTGAAAACTGCATCTGAAAGCAACTTTATGTTCAATTATCTTAATGATGG 240
Db 193 TTATAATTGAAAACTGCATCTGAAAGCAACTTTATGTTCAATTATCTTAATGATGG 252
QY 241 TGTTTATGACTAATACACTGATTTTTCATAAGAAACCACGTTAAATAATTTTAT 300
Db 253 TGTTTATGACTAATACACTGATTTTTCATAAGAAACCACGTTAAATAATTTTAT 312
QY 301 TTTAAAAATAAGCCTGTGTTCAGCTCTGATCATATTCTTTTATTTGATTTGGGAAGA 360
Db 313 TTTAAAAATAAGCCTGTGTTCAGCTCTGATCATATTCTTTTATTTGATTTGGGAAGA 372
QY 361 AAATACCTGTTCTGATGATGACATGAAATGCAAAATTTTAGATTTTAACTCNCATAATTT 420
Db 373 AAATACCTGTTCTGATGATGACATGAAATGCAAAATTTTAGATTTTAACTCNCATAATTT 432
QY 421 AAGAACTATTGAGAAATTGATTAAATGACATGAAGTGACACAACACTAACTGCG 474
Db 433 AAGAACTATTGAGAAATTGATTAAATGACATGAAGTGACACAACACTAACTGCG 486
RESULT 5
AAAX40097
ID AAAX40097 standard; DNA; 843 BP.
XX
AC AAAX40097;
XX
DT 02-JUL-1999 (first entry)
XX
DE Gastric cancer associated gene.
XX
KM Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9904265-A2.
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98WO-US14679.
XX
PR 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfeundschnuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
PS Claim 67; Page 696; 787bp; English.
XX

CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.

XX Sequence 843 BP; 237 A; 184 C; 89 G; 215 T; 118 other;

Query Match 59.9%; Score 411.8; DB 20; Length 843;
Best Local Similarity 71.8%; Pred. No. 8.8e-62;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

OY 1 ATTTAATTTCTATTAAACATCTCTCAAGCATTTATCTCTATATCTCACTGAA 60
DB 15 ATTTAATTTCTATTAAACATCTCTCAAGCATTTATCTCTATATCTCACTGAA 74
OY 61 TTTTAAANAATACATTAAGTATAGAAAAAAGAGAAAAAGATNAATGAGATAATTAA 120
DB 75 TTTNANAANAATACCTTTNNTNTTAAAAAACCTGGAAGAAAAAATAATNGCAATANTTAA 134
OY 121 ACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATAAATTGAATGAGA 180
DB 135 CTTNCTTGAAAAANGAATTTNTACCAANGAAGAAANCNTNTAATNGAANTNAAA 194
OY 181 TTATAATTTGAAAACTGCATCTGAAAGCAAACTTTATGTTCAATATCTTAATGATG 240
DB 195 TTATANTTNGAAANCGGCGNNCGAAACCAANTTNATGTCCTCAATATCTTAANGAGGG 254
OY 241 TGTATTATGACTAATACACTGATTTTCAATTAAGAAAAACCATGTTAAAAATTTTAT 300
DB 255 NNTTANNACTAATNCCNGATTTTCCAATNNGAANCCNNNTTAAANNTTTTAT 314
OY 301 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCAATTTCTTTATTTGATTTGGGAAGA 360
DB 315 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCAATTTCTTTATTTGATTTGGGAAGA 374
OY 361 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTATCTCNCATATTTT 420
DB 375 AAATNCGTCCNNAATACCNNGAANNGCAANTTTTAAATTTTAAACCCCTANTTTT 434
OY 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGACACAACTAATTTAGGCCAGCT 480
DB 435 AAAANCTATNGAAAAANTNGATTANNGACTGAAATTGC-CAACCTANTNCGGCCAACCN 493
OY 481 GTTGGCATTTGTTTCTTACTTAGTTCTCCCAAGAAAACTTTAAACTGAACTCTCAGC 540
DB 494 GTGGCCTNTGTTTCTTACTTANTTCCCCCAAGAAANNCTTAANCNGAANCTCCNCC 553
OY 541 NGATTAACCTTAATATATCTTTGTTAGCCAAAC--AAAACTTTTGTGTTACATAGTTCT 598
DB 554 AAAATAACCTTAATATATCTTTGTTAGCCAAACAAACCTTTTNGTTACNTANTCTCT 613
OY 599 TTGATTTTACTGTTCTTAATTTTATTTCTGAACCTCATTTTCCCGACAGCATATATAC 658
DB 614 TGGGATTTAAAGGTCGCCCAATTTNATCCNGAACCAANTTTTCCCCCAACCATANTTAC 673
OY 659 CCTATTAACTTTGTTATGACAGAGTTGT 687
DB 674 CAT-TTACCTTGTTAAGGCNCAGTNGTT 701

RESULT 6
AAT19975/c
ID AAT19975 standard; cDNA to mRNA; 351 BP.

XX AAT19975;
AC 17-JUL-1996 (first entry)
DT 17-JUL-1996 (first entry)
XX Human gene signature HUMGS01107.
DE Human gene signature HUMGS01107.
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX Homo sapiens.
XX WO9514772-A1.
XX PN 01-JUN-1995.
XX PD 11-NOV-1994; 94WO-JP01916.
XX PF 12-NOV-1993; 93JP-0355504.
XX PR (MATS/) MATSUBARA K.
XX PA (OKUB/) OKUBO K.
XX PI Matsubara K, Okubo K;
XX DR WPI; 1995-206931/27.
XX DR WPI; 1995-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
PS Claim 1; Page 525; 2245pp; Japanese.
XX
XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 351 BP; 117 A; 44 C; 44 G; 141 T; 5 other;

Query Match 43.8%; Score 300.8; DB 16; Length 351;
Best Local Similarity 97.0%; Pred. No. 7.8e-43;
Matches 324; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

OY 1 ATTTAATTTCTATTAAACATCTCTCAAGCATTTATTTAT-CTATATCTCACTGA 59
DB 334 ATTTAATTTCTATTAAACATCTCTCAAGCATTTATTTATNCCCTATATCTCACTGA 275
OY 60 ATTTAANAATAACATTAAGTATAGAAAAAAGAGAAAAAGATNAATGAGATAATTA 119
DB 274 ATTTAAGAATAACATTAAGTATAGAAAAAAGAGAAAAAGATNAATGAGATAATTA 215
OY 120 AACTTACATGAAAAAGAAAAATTATACAAAGAGACTGAGAACGTTATTAATTGAATGAG 179
DB 214 AACTTACATGAAAAAGAAAAATTATACAAAGAGACTGAGAACGTTATTAATTGAATGAG 155
OY 180 ATTATAATTTGAAAAAAGTGCATCTGAAAGCAAACTTTATGTTCAATTAATTTAATGATG 239
DB 154 ATTATAATTTGAAAAAAGTGCATCTGAAAGCAAACTTTATGTTCAATTAATTTAATGATG 95

```
QY      240 GTCGTTTATGACTAATAACACTGATTTTCAATAAGGAACCCTGTAAATAATT-TT 298
        |||
Db       94 GTCGTTTATGACTAATAACACTGATTTTCAATAAGGAACCCTGTAAATAATTNTN 35
        |||
QY      299 ATTTTAAAAATAAGCCTGTGTTCAAGCTTGATC 332
        |||
Db       34 NTTTTAAAAATAAGCCTGTGTTCAAGCTTGATC 1
```

RESULT 7
AAH57293/c
ID AAH57293 standard; cDNA; 262 BP.

AC AAH57293;

DT 10-SEP-2001 (first entry)

DE Human liver specific cDNA sequence SEQ ID NO:133.

KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
 KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

OS Homo sapiens.

PN W0200132927-A2.

PD 10-MAY-2001.

PF 02-NOV-2000; 2000WO-US30396.

PR 04-NOV-1999; 99US-0163508.

PA (INCY-) INCYTE GENOMICS INC.

PI Sornasse T, Seilhamer JJ, Watson GA;

DR WPI; 2001-291057/30.

PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
PT

PS Claim 1; Page 117; 327pp; English.

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide sequences (I). (I) can have cytostatic, immunomodulatory and neuroprotective activities, and can be used in gene therapy. (I) and proteins (II) encoded by them are used in high throughput screening assays to select DNA molecules, RNA molecules, peptide nucleic acids, mimetics, peptides, proteins, agonists, antagonists, antibodies or their fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents. Expression of (I) in a sample indicates the differentiation of embryonic stem cells into a tissue selected from brain, heart, kidney, liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used to produce an expression profile that defines a metabolic or developmental process, treatment, condition, disease or disorder. The gene profile can be used for diagnosis, prognosis or monitoring of treatments and for investigating a predisposition to a disorder where the gene is associated with a cancer, immunopathology or neuropathology.

Sequence 262 BP; 103 A; 38 C; 34 G; 87 T; 0 other;

Query Match	38.0%	Score 261;	DB 22;	Length 262;
Best Local Similarity	99.6%;	Pred. No. 4.8e-36;		
Matches 261; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	243	TTTTATGCTAATACACTGATTTTCAATAAGGAACCCAGTTAAAAATATTTTATTTT	302
Db	262	TTTTATGCTAATACACTGATTTTCAATAAGGAACCCAGTTAAAAATATTTTATTTT	203

[illegible]

Qy	363	ATACTGTTTCTGATAGCATGAAATGCAAAATTTTAGATTTTAACTCNCATAATTTTAA	4222
Db	142	ATACTGTTTCTGATAGCATGAAATGCAAAATTTTAGATTTTAACTCNCATAATTTTAA	83

423 GAACTATTGAGAAATTGATTATATGACATGAAGTCACACACTAATTACTGGCCAGCTGT 482

Db 82 GAACTATTGAGAAATTGATTATGACATGAAGTCACACACTAATTACTGGCCAGCTGT 23

QY 483 TGGCATTGTGTTCTTACTTAG 504

Db 22 TGGCATTGTGTTCTTACTTAG 1

RESULT 8
ABX83093/c

AC ABX83093;

DT 24-APR-2003 (first entry)

DE Corn ear-derived polynucleotide (cpd) #1553.

KM Corn ear-derived polynucleotide; cdp; cdna library; SATMON022;
KM SATMON023; structural gene; functional gene; regulatory gene;
KM corn ear-specific profile; gene transcription; gene expression;
KM hybrid plant; desirable trait expression; plant breeding program;
KM inheritance; desired characteristic; growth; development;
KM disease resistance; environmental adaptability; quality; yield;
KM multigene trait; plant; gene; ss.

OS Zea mays.

PN US6476212-B1.

PD 05-NOV-2002.

PF 14-MAY-1999; 99US-0313294.

PR 26-MAY-1998; 98US-086722P.

PA (INCY-) INCYTE GENOMICS INC.

PI Lalgudi RV, Ito LY, Sherman BK;

DR WPI; 2003-208840/20.

PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with
PT growth, development -

PS Examples; SEQ ID No 1553; 390pp; English.

The present invention relates to the isolation of corn ear-derived polynucleotides (cdps) from the corn (*Zea mays*) cDNA libraries SATMON022 and SATMON023. Some of the cpds uniquely identify structural, functional, and regulatory genes of corn ear. The polynucleotides sequences are useful for detecting cpds in a sample, for producing a corn ear-specific profile of gene transcription, for detecting altered gene expression in inbred or hybrid plants, and for screening several molecules for specific binding to the polynucleotide. The cpds are useful to identify, isolate, or extend identical or related corn-ear nucleic acid sequences from DNA libraries, and in nucleic acid amplification or hybridisation techniques to follow the expression of desirable traits through plant breeding programs. Preferably, the cpds are used to identify, evaluate, alter, or follow the inheritance of desired characteristics associated with growth and development, disease resistance, environmental adaptability,

CC quality, and yield of corn. The cdps are also useful as molecular
CC markers for studying inheritance and multigene traits in a plant
CC breeding program. The cdps are useful for producing purified corn-ear
CC polypeptides by recombinant techniques. They are also useful in
CC diagnostic assays to detect or confirm conditions or diseases
CC associated with abnormal levels of cdp expression. ABX81541-ABX89140
CC represent corn ear-derived polynucleotides (cpds) of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/psipdidentry.html.

XX SQ Sequence 173 BP; 54 A; 20 C; 19 G; 80 T; 0 other;

Query Match 23.4%; Score 161; DB 25; Length 173;
Best Local Similarity 98.8%; Pred. No. 5.7e-19;
Matches 161; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 TTTAATTTCTATTAACATTCTTCTCAAGCATTATTTATCCTATATCTCACTGAATT 62

DB 163 TTTAATTTCTATTAACATTCTTCTCAAGCATTATTTATCCTATATCTCACTGAATT 104

OY 63 TTAANAAATTAACATTAGTATTAGAAAACTAGAAAAAGATNAATGCAGATAATTAAAC 122

DB 103 TTAAGAAATTAACATTAGTATTAGAAAACTAGAAAAAGATNAATGCAGATAATTAAAC 44

OY 123 TTACATGAAAAAGAAAAATTATACAAAGGACTGAGAACGTTA 165

DB 43 TTACATGAAAAAGAAAAATTATACAAAGGACTGAGAACGTTA 1

RESULT 9

AAQ76489/C
ID AAQ76489 standard; DNA; 263 BP.

AC AAQ76489;

DT 25-MAR-2003 (updated)

DT 23-SEP-1994 (first entry)

DE Human genome fragment.

KW Brain; placenta; bone marrow; genetic analysis; gene mapping;

KW detection; homology; human; adrenal tissue; ds.

OS Homo sapiens.

PN WO9401548-A2.

PD 20-JAN-1994.

PF 13-JUL-1993; 93WO-GB01467.

PR 13-JUL-1992; 92GB-0014857.

PA (MED1-) MEDICAL RES COUNCIL.

PI Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;

PI Sibson DR, Starkey M;

DR WPI; 1994-035056/04.

XX New nucleic acid fragment encoding gene products - can be used

XX PT for genetic analysis and mapping

XX PS Claim 1; Page 64; 616pp; English.

CC Human nucleic acid fragments, isolated from brain adrenal tissue,

CC the placenta or bone marrow comprise any of: (A) a sequence
CC selected from (AAQ76401-Q77613), (B) an allelic variation of a
CC sequence as described in (A), or (C) a sequence complementary
CC to (A) or (B).
CC (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 263 BP; 97 A; 38 C; 45 G; 80 T; 3 other;

Query Match 18.4%; Score 126.2; DB 15; Length 263;
Best Local Similarity 81.5%; Pred. No. 4.9e-13;
Matches 194; Conservative 0; Mismatches 36; Indels 8; Gaps 4;

OY 450 TGAAGTGACACACTAATTACTGGCCAGCTGTGGCATTTGTTCTTACTAGTTCTC 509

DB 263 TGGAGTACACAGCACTGATTACTGGCCAGCTGTGGCATTTGTTCTTACTAGTTCTC 205

OY 510 CCAAGGAAACTCTTAACTGATCTTCACGNGAATAACCTTAATATACTTTGTAGCC 569

DB 204 ACGAGGAAACTCTTAACTGATCTT--AGCAGAAAATCCTTAATATACTTTGTAGCAA 147

OY 570 AAACAAACTTTTGTGTTTACATAGTTCTTTGATTTTACTGTTCTTAATTTTACTGCA 629

DB 146 AACAAAGCTTTTGTGTTTACATCGTTCTTTGATTTTATTTGTTTAAATTTTACTGCA 87

OY 630 AACTCCATTTTTC-CCCAGACCATTAATCCCTATTTAACCTTTGTTATGCAGAGTTGT 686

DB 86 AACTGATTTTACNNCCAAACCATTAATTC----TATTAATTTGTTATGCAGAGTTGT 33

RESULT 10

ABX40797
ID ABX40797 standard; cDNA; 397 BP.

AC ABX40797;

DT 20-FEB-2003 (first entry)

DE Bovine EST associated with lactation/muscle/fat deposition #5962.

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

KW muscle deposition; fat deposition; genome mapping; gene identification;

OS Bos Taurus.

PN US2002137139-A1.

PD 26-SEP-2002.

PF 24-SEP-2001; 2001US-0960352.

PR 12-JAN-1999; 99US-115707P.

PR 11-JAN-2000; 2000US-0480902.

PA (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

PI Byatt JC, Mathialagan N, Tao N, Warren WC;

DR WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat

XX PT deposition, useful for genome mapping, gene identification and

XX PT analysis, cattle breeding, or for genetically improving cattle

XX PS Claim 2; SEQ ID No 5962; 245pp; English.

CC The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMFD), derived

CC from cattle, and the LMFD nucleic acid can specifically hybridise to a

CC second nucleic acid molecule comprising any of 15112 nucleotide

CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a

CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docID=20020137139.

XX SQ Sequence 397 BP; 144 A; 75 C; 45 G; 133 T; 0 other;

Query Match 15.1%; Score 103.6; DB 25; Length 397;
Best Local Similarity 71.9%; Pred. No. 3.5e-09;
Matches 192; Conservative 0; Mismatches 65; Indels 10; Gaps 4;

OY 421 AAGACTATTGAGAATGATTAATGACATGAGTGCACACACTAATTACTGGCCAGCT 480
Db 14 AATTATCACTGAGAAATATATTAAGTCAATTAATGCACACATTAATTACTGGTCACT 73
OY 481 GTTGCCATTGTGTTCTTACTTAAGTCTCTCCCAAGGAAAACTTAACTGAATCTTCAGC 540
Db 74 AATAATGATAT-TCTGTTCTTCAATCTCCCAAGGAAAACT-----TGAATTTGGGC 125
OY 541 NGATAAACCTTAATATATCTTTGTTAGCCCAACAACTTTTGTTCATAGTCTTT 600
Db 126 AGAATAATCTTCAAATATACATTATTAGCAAAATGAGAGCTTCTGTACATACCT-TTT 184
OY 601 GGATTTTACTGTTCCCTAATTTTATCTGAAACTCCATTTTCCCGAGACCATTAATACC 660
Db 185 GTATTTGCTATTTCTAATCTTAATCTTAAACTCAATTTTACCCCAACCATGATTACCA 244
OY 661 TATTTAATCTTTGTATGCACAGTTGTT 687
Db 245 TA-TTAATCTTTGTAATGCACAGTTGTT 270

RESULT 11

ABI99762/c
ID ABI99762 standard; cDNA; 2152 BP.

XX AC ABI99762;

DT 07-MAR-2002 (first entry)

XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:842.

KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KM vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

XX PN WO200188188-A2.

XX PD 22-NOV-2001.

XX PF 18-MAY-2001; 2001WO-JP04192.

XX PR 18-MAY-2000; 2000JP-0145977.

XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX DR WPI; 2002-034733/04.

DR P-PSDB; ABB57301.

XX PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -

XX PS Claim 2; Page 2087-2091; 2690pp; English.

XX CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX SQ Sequence 2152 BP; 586 A; 455 C; 576 G; 535 T; 0 other;

Query Match 12.5%; Score 85.8; DB 24; Length 2152;
Best Local Similarity 65.2%; Pred. No. 3.8e-06;
Matches 172; Conservative 0; Mismatches 88; Indels 4; Gaps 3;

OY 421 AAGACTATTGAGAAATGATTAATGACATGAGTGCACACACTAATTACTGGCCAGCT 480
Db 2130 AACTATTAGTGAAGAACTGATGACAGGTAATGAGTGTGAGGCACATAATCTGTCACCT 2071
OY 481 GTTGCCATTGTGTTCTTACTTAAGTCTCTCCCAAGGAAAACTCTTAACTGAATCTTCAGC 540
Db 2070 -ATGGAATCTGTTCTTCTCTAGTCCCCCAAGAGAAATGTACACCGAGTGTCCAAC 2012
OY 541 NGATAACCTTAATATATACCTTTGTTAGCCAAACAACTTTTGTTCATAGTCTTT 600
Db 2011 ACAGTGATCCTCAATATACTTCTTAAGCAAAAC--AAGGCTTTGTTAATCTAGTCTTCA 1954
OY 601 GGATTTTACTGTTCCCTAATTTTATCTGAACTCCA-TTTTCCCGACCATTAATTACC 659
Db 1953 GATTTTACTAGTTCTGAATTTTATTTCTGAAACTCAAGTTTAAACCCCAACCATCATTTGAC 1894
OY 660 CTATTTAATCTTTGTATGCACAGT 683
Db 1893 ATAGTCAGCTTCATCAAGCACAGT 1870

RESULT 12

ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.

XX AC ABZ10246;

DT 16-JAN-2003 (first entry)

XX DE Haematopoietic cell proliferation disorder related DNA sequence #386.

KM Human; haematopoietic cell proliferation disorder; cytostatic;
KM gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KM cytosine methylation state; gene; ds.

OS Homo sapiens.

XX PN WO200277272-A2.

XX PD 03-OCT-2002.

XX PF 26-MAR-2002; 2002WO-EP03401.

PR 26-MAR-2001; 2001US-278333P.
XX (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
PI Pelet C, Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent
PT that distinguishes between methylated and non-methylated CpG
PT dinucleotides -
XX
XX
PS Claim 28; SEQ ID 386; 117bp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between hematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. AB209861 to AB211118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy hematopoietic cells and proliferative
CC disorder hematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of hematopoietic cell proliferation disorder related
CC DNA sequences. The nucleotide sequences from the present invention can
CC also be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC hematopoietic cell proliferative disorders. The present method enables
CC a highly specific classification of hematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients.
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 other;

Query Match 10.4%; Score 71.6; DB 25; Length 8056;
Best Local Similarity 50.3%; Pred. No. 0.00098;
Matches 224; Conservative 0; Mismatches 217; Indels 4; Gaps 2;

QY 2 TTTTATTTCTATTAACATCTCTCAAGACATTATTTTCTATATCTCACTGAAT 61
DB 2056 TTAATAAATTTTATTAATAATTTTAAATATTAATAATTAATTAATTTAAT 2115
QY 62 TTTAANAATAACATTAAGTATTGAAAACTAGAAAAAGATNAATGCAGTAATTAA 121
DB 2116 ATAAATTTAATAATTTTATTTTATTAATAATAATAATAATAATAATAATAATA 2175
QY 122 CTACATGAAAAAGAAATTAT-AACAAGACTGAGAACGTTATAATTGAATGAG 179
DB 2176 TTTTATTAATAAATTTTAAATAATAATAATAATAATAATAATAATAATAATT 2235
QY 180 ATTATAATTTGAAACCTGCATCTGAAAGCAACTTATTGTTCAATTATCTTAATGATG 239
DB 2236 ATAAAAATTTATTTTAAATAATTAAATAATAATAATAATAATAATAATAATT 2295
QY 240 GTGTTTATGACTAATACACTGATTTTTCATAGAGAACCCATGTTAAATAATTTT 299
DB 2296 TTATTTTAAATAATTAATTAATTTT--GTAATTAATAATAATTTATTAATTTTAA 2353
QY 300 TTTTAAAAATAGCCGTGTCAAGCTCGATCAATTTCTTTATTTTGAATTTGGGAG 359
DB 2354 ATATAATTTTATTTTATTAATAATTAATTTATTAATTTTGAATAATAATAA 2413
QY 360 AAAATACCTGTTCTGATAGCATGAATGCAAAATTTTGAATTTTAACTCNCCTAATT 419
|||||

DB 2414 AAAAAATAATTTATTTATTTTAAAAATAATTTTATTTTATTTATATATTTT 2473
QY 420 TAAGAACTATTGAGAAATTGATTAA 444
DB 2474 TATTAATTTGTTTAAATAATAATTAA 2498

RESULT 13
AB210246/C
ID AB210246 standard; DNA; 8056 BP.
XX
XX AC AB210246;
XX
DT 16-JAN-2003 (first entry)
XX
DE Hematopoietic cell proliferation disorder related DNA sequence #386.
XX
XX Human; hematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP03401.
XX
PR 26-MAR-2001; 2001US-278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T;
PI Pelet C, Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent
PT that distinguishes between methylated and non-methylated CpG
PT dinucleotides -
XX
PS Claim 28; SEQ ID 386; 117bp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between hematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. AB209861 to AB211118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy hematopoietic cells and proliferative
CC disorder hematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of hematopoietic cell proliferation disorder related
CC DNA sequences. The nucleotide sequences from the present invention can
CC also be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC hematopoietic cell proliferative disorders. The present method enables
CC a highly specific classification of hematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients.
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 other;

Query Match 10.2%; Score 70; DB 25; Length 8056;

ID ABL34598 standard; DNA; 7676 BP.
XX ABL34598;
AC
XX
DT 26-MAR-2002 (first entry)
XX
DE Human metastasis associated gene SEQ ID NO: 151.
XX
KM Metastasis associated gene; cytostatic; gene therapy; cancer;
XX cytosine methylation; gene; db.
OS Homo sapiens.
XX
PN WO200177376-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-EP03970.
XX
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-010922/01.
XX
PT New nucleic acid derived from chemically treated metastasis genes,
PT useful for diagnosis of cancers by analysis of cytosine methylation,
PT also for treatment -
XX
PS Claim 1; SEQ ID NO 151; 23bp + Sequence Listing; English.
XX
CC The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention.
XX
SQ Sequence 7676 BP; 2396 A; 45 C; 1348 G; 3887 T; 0 other;

Query Match 9.7%; Score 66.8; DB 24; Length 7676;
Best Local Similarity 48.5%; Pred. No. 0.0065;
Matches 210; Conservative 0; Mismatches 220; Indels 3; Gaps 1;

QY 21 ATTCTCTCAAGCATTTTATCTATATCTCAGTAATTTAANAATACATT--- 77
DB 3697 ATTTGTTAATAATATATATATATTAGTGTAGTAAATTTAAATTGTTAAATATTAA 3756
QY 78 AGTATTAGAAAACTAGAAAAAGATNAATGACAGATAATTAACTTACATGAAAAAGGA 137
DB 3757 TATATTAAATTATATATAATAAATAGATATGAAATATTAGAAATGAAAGTTT 3816
QY 138 AAATTATACAAAGACTGAGACGTTATAATTGAAATGAGATTATAATTGAAAACTG 197
DB 3817 AATGTATAGTTAAATTTTATATAAGGAATATTTTATATAGATATAAGAAATTT 3876
QY 198 CATCTGAAGCAACTTATTTGTTCAATATCTTAATGATGCTTTTATGACTAATAC 257
DB 3877 AATTTAAATTAAAAATTTAGTTTTTTTATGATAAAGTTTTTTTATATTATAATA 3936
QY 258 ACTGATTTTCAATAAGGAAACCATGTTAAAAATATTTTATTTAAAAATAAGCCTGT 317
DB 3937 GATTAATTTTTTTTTTTTAAAGTAGAGTTTATAAATTTTTTTAGTGAGTTATTAAGA 3996
QY 318 GTTCAAGCTCTGATCATATTTCTTTATTTGATTTGGGAGAAAAATACTGTTCTGATA 377
DB 3997 TTTTGTGTGTGTTTTTTTATTTGAATATTTTGTGAAATTAATTAATTAGTTTTA 4056
QY 378 GCATGAATGCAAAATTTTATAGATTTTAACTCNCCTAATTTTAAGAACTATTGAGAAAT 437

DB 4057 AAATAGTAATTAATTAATTTTTTTATTTTAAATTAGTAATAAGTAATAATTAGAGATT 4116
QY 438 TGATTAATGACAT 450
DB 4117 AGTAGAATAGTAT 4129

Search completed: November 27, 2003, 06:23:43
Job time : 226.747 sec8

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 06:06:02 ; Search time 259.057 Seconds
(without alignments)
8737.270 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 687

Sequence: 1 attctaattctctataaac.....cttcttatgcacagttgtc 687

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	683	99.4	687	9	US-09-835-992A-20
2	638.4	92.9	1052	10	US-09-880-107-3296
3	611.8	89.1	714	9	US-09-835-992A-19
4	471	68.6	506	10	US-09-796-692-7228
5	471	68.6	506	14	US-10-040-862-7228
6	467.8	68.1	486	10	US-09-998-598-578
7	411.8	59.9	843	9	US-09-835-992A-22
8	161.8	23.6	510	14	US-10-102-524-161
9	161.8	23.6	510	14	US-10-102-524-188
10	161.8	23.6	510	14	US-10-102-524-645
11	103.6	15.1	397	10	US-09-960-352-5962
12	66.8	9.7	7676	12	US-10-240-485-151
13	65	9.5	113515	12	US-10-311-455-2147
14	63.4	9.2	960	14	US-10-198-846-6381
15	63	9.2	6106	12	US-10-311-455-1445
16	62	9.0	3673778	12	US-10-312-841-1

C 17	61.8	9.0	291	10	US-09-960-352-13694	Sequence 13694, A
C 18	60.4	8.8	6944	14	US-10-172-086-111	Sequence 111, App
C 19	60.2	8.8	6509	12	US-10-311-455-199	Sequence 199, App
C 20	60.2	8.8	6912	12	US-10-240-453-245	Sequence 245, App
C 21	59.8	8.7	17721	12	US-10-311-455-1702	Sequence 1702, App
C 22	59.6	8.7	10467	12	US-10-240-453-327	Sequence 327, App
C 23	59.4	8.6	3991	14	US-10-074-045-60	Sequence 60, App1
C 24	59.2	8.6	640681	10	US-09-790-988-1	Sequence 1, App1
C 25	59	8.6	3673778	12	US-10-312-841-2	Sequence 2, App1
C 26	58.8	8.6	449	10	US-09-960-352-13260	Sequence 13260, A
C 27	58.4	8.5	424	10	US-09-960-352-11218	Sequence 11218, A
C 28	58.2	8.5	5962	12	US-10-311-455-1260	Sequence 1260, App
C 29	58	8.4	18218	12	US-10-311-455-1922	Sequence 1922, App
C 30	57.8	8.4	9219	12	US-10-311-455-2400	Sequence 2400, App
C 31	57.8	8.4	3673778	12	US-10-312-841-1	Sequence 1, App1
C 32	57.6	8.4	9810	12	US-10-311-455-400	Sequence 400, App
C 33	57.6	8.4	3673778	12	US-10-312-841-2	Sequence 2, App1
C 34	57.2	8.3	640681	10	US-09-790-988-1	Sequence 1, App1
C 35	57	8.3	480	10	US-09-960-352-5301	Sequence 5301, App
C 36	57	8.3	13712	12	US-10-311-455-1504	Sequence 1504, App
C 37	56.8	8.3	11745	12	US-10-240-453-206	Sequence 206, App
C 38	56.8	8.3	18624	12	US-10-311-455-1675	Sequence 1675, App
C 39	56.2	8.2	416	10	US-09-960-352-4584	Sequence 4584, App
C 40	56.2	8.2	8979	12	US-10-311-455-758	Sequence 758, App
C 41	56.2	8.2	12138	12	US-10-311-455-1916	Sequence 1916, App
C 42	56.2	8.2	12138	12	US-10-240-453-210	Sequence 210, App
C 43	55.8	8.1	5276	12	US-10-311-455-124	Sequence 124, App
C 44	55.8	8.1	7049	12	US-10-311-455-130	Sequence 130, App
C 45	55.8	8.1	7049	12	US-10-240-452-6	Sequence 6, App1

ALIGNMENTS

RESULT 1

US-09-835-992A-20

Sequence 20, Application US/09835992A

Patent No. US20020037541A1

GENERAL INFORMATION:

APPLICANT: Obata, Yuichi

TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AND

TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING GASTRIC CANCER

FILE REFERENCE: L0461/7112

CURRENT APPLICATION NUMBER: US/09/835, 992A

CURRENT FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: US 08/896, 164

PRIOR FILING DATE: 1997-07-17

NUMBER OF SEQ ID NOS: 87

SOFTWARE: PatentIn version 3.0

SEQ ID NO 20

LENGTH: 687

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: Unsure

LOCATION: (67)..(67)

OTHER INFORMATION: n = a, c, g or t

NAME/KEY: Unsure

LOCATION: (105)..(105)

OTHER INFORMATION: n = a, c, g or t

NAME/KEY: Unsure

LOCATION: (412)..(412)

OTHER INFORMATION: n = a, c, g or t

NAME/KEY: Unsure

LOCATION: (541)..(541)

OTHER INFORMATION: n = a, c, g or t

US-09-835-992A-20

Query Match 99.4%; Score 683; DB 9; Length 687;

Best Local Similarity 100.0%; Pred. No. 5.8e-114;

Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATTTTAATTTCTATTAACATCTCTCAAGCATTATTTATCTATATCTCACTGAA 60

```
|||||
Db 1 ATTTTAATTTTCTATTAACATCTTCTCAAGCATTAATTTATCCTATATCTCACTGAA 60
QY 61 TTTTAANAATAACATTAGTATTTAGAAAAAAGATNAATGCAGTAATTAA 120
Db 61 TTTTAANAATAACATTAGTATTTAGAAAAAAGATNAATGCAGTAATTAA 120
QY 121 ACTTACATGAAAAAGAAAAATTATAACAAGAGCTGAGAACGTTATAAATGAAATGAGA 180
Db 121 ACTTACATGAAAAAGAAAAATTATAACAAGAGCTGAGAACGTTATAAATGAAATGAGA 180
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATTATCTTAATGATGG 240
Db 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATTATCTTAATGATGG 240
QY 241 TGTATTATGACTAATACACTGATTTTCAATAAGGAAACCCATGTTAAAAATATTTTAT 300
Db 241 TGTATTATGACTAATACACTGATTTTCAATAAGGAAACCCATGTTAAAAATATTTTAT 300
QY 301 TTTAAAAATAAGCCCTGTGTCTCAAGCTCTGATCATATTTCTTTATTTGATTTGGGAAGA 360
Db 301 TTTAAAAATAAGCCCTGTGTCTCAAGCTCTGATCATATTTCTTTATTTGATTTGGGAAGA 360
QY 361 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 420
Db 361 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 420
QY 421 AAGAACTATGAGAAATGATTAATGACATGACACAACACTAATTAATCTGACCAGCT 480
Db 421 AAGAACTATGAGAAATGATTAATGACATGACACAACACTAATTAATCTGACCAGCT 480
QY 481 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC 540
Db 481 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC 540
QY 541 NGATAACCTTAATAATACTTTGTTAGCCAAACAACCTTTTGTGTTACATAGTTCTTT 600
Db 541 NGATAACCTTAATAATACTTTGTTAGCCAAACAACCTTTTGTGTTACATAGTTCTTT 600
QY 601 GGATTTTACTGTTCTTAATTTTATCTGAAACCTCAATTTTCCCGACAGCATTAATTAACC 660
Db 601 GGATTTTACTGTTCTTAATTTTATCTGAAACCTCAATTTTCCCGACAGCATTAATTAACC 660
QY 661 TATTTAACTTTGTTATGACAGTTGTT 687
Db 661 TATTTAACTTTGTTATGACAGTTGTT 687

RESULT 2
US-09-880-107-3296/c
; Sequence 3296, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3296
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U11313
```

```
US-09-880-107-3296
Query Match 92.9%; Score 638.4; DB 10; Length 1052;
Best Local Similarity 98.1%; Pred. No. 7.4e-106;
Matches 675; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

QY 1 ATTTTAATTTTCTATTAACATCTTCTCAAGCATTAATTTATCCTATATCTCACTGAA 60
Db 1014 ATTTTAATTTTCTATTAACATCTTCTCAAGCATTAATTTATCCTATATCTCACTGAA 955
QY 61 TTTTAANAATAACATTAGTATTTAGAAAAAAGATNAATGCAGTAATTAA 120
Db 954 TTTTAAGAAATAACATTAGTATTTAGAAAAAAGATNAATGCAGTAATTAA 895
QY 121 ACTTACATGAAAAAGAAAAATTATAACAAGAGCTGAGAACGTTATAAATGAAATGAGA 180
Db 894 ACTTACATGAAAAAGAAAAATTATAACAAGAGCTGAGAACGTTATAAATGAAATGAGA 835
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATTATCTTAATGATGG 240
Db 834 TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATTATCTTAATGATGG 775
QY 241 TGTATTATGACTAATACACTGATTTTCAATAAGGAAACCCATGTTAAAAATATTTTAT 300
Db 774 TGTATTATGACTAATACACTGATTTTCAATAAGGAAACCCATGTTAAAAATATTTTAT 715
QY 301 TTTAAAAATAAGCCCTGTGTCTCAAGCTCTGATCATATTTCTTTAATTTGATTTGGGAAGA 360
Db 714 TTTAAAAATAAGCCCTGTGTCTCAAGCTCTGATCATATTTCTTTAATTTGATTTGGGAAGA 655
QY 361 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 420
Db 654 AAATACGTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 595
QY 421 AAGAACTATGAGAAATGATTAATGACATGAGTGACACAACACTAATTAATCTGACCAGCT 480
Db 594 AAGAACTATGAGAAATGATTAATGACATGAGTGACACAACACTAATTAATCTGACCAGCT 535
QY 481 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC 540
Db 534 GTTGCAATGTGTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC 475
QY 541 NGATAA-CCTTAATATACTTTGTTAGCCAAACAACAA-CCTTTTGTGTTACATAGTTCT 598
Db 474 AGAATAATCTTAATATATCTTTGTTAGCCAAACAACAAAGCTTTTGTGTTACATAGTTCT 415
QY 599 TTGATTTTACTGTTCTTAATTTTATCTGAAACCTCAATTTTCCCGACAGCATTAATTAACC 658
Db 414 TTGATTTTACTGTTCTTAATTTTATCTGAAACCTCAATTTTCCCGACAGCATTAATTAACC 355
QY 659 CCTATTTAACTTTGTTATGACAGTTGT 686
Db 354 CATATTTAACTTTGTTATGACAGTTGT 328

RESULT 3
US-09-835-992A-19
; Sequence 19, Application US/09835992A
; Patent No. US20020037541A1
; GENERAL INFORMATION:
; APPLICANT: Obata, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AND
; FILE REFERENCE: L0461/7112
; CURRENT APPLICATION NUMBER: US/09/835,992A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 08/896,164
; PRIOR FILING DATE: 1997-07-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 714
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unscore
; LOCATION: (243)..(243)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unscore
; LOCATION: (373)..(373)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unscore
; LOCATION: (437)..(437)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unscore
; LOCATION: (702)..(702)
; OTHER INFORMATION: n = a, c, g or t
US-09-835-992A-19
```

```

Query Match      89.1%; Score 611.8; DB 9; Length 714;
Best Local Similarity 97.1%; Pred. No. 4.1e-101;
Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;
```

```

QY 1 ATTTAATTTTCTATTAACATCTCTCAAGCATTATTTTACCTATATCTCACTGAA 60
    |||||||
DB 15 ATTTAATTTTCTATTAACATCTCTCAAGCATTATTTTACCTATATCTCACTGAA 74
QY 61 TTTTAAATAATAACATTAGTATTTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 120
    |||||||
DB 75 TTTTAAATAATAACATTAGTATTTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 134
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGACTGAGAACGTTATTAATTGAATGAGA 180
    |||||||
DB 135 ACTTACATGAAAAAGAAAAATTATACAAAGACTGAGAACGTTATTAATTGAATGAGA 194
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 240
    |||||||
DB 195 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 254
QY 241 TGTTTATGACTAATACATGATTTTCAATAAGGAACCCAGTTAAAAATATTTTAT 300
    |||||||
DB 255 TGTTTATGACTAATACATGATTTTCAAGAAGGAACCCAGTTAAAAATATTTTAT 314
QY 301 TTTTAAATAAAGCCTGTGTCAAGCTCTGATCATATTTCTTTAATTTTGATTTGGGAAGA 360
    |||||||
DB 315 TTTTAAATAAAGCCTGTGTCAAGCTCTGATCATATTTCTTTAATTTTGATTTGGGAAGA 374
QY 361 AAATACGTCTTCTGATGATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 420
    |||||||
DB 375 AAATACGTCTTCTGATGATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 434
QY 421 AAGAACTATTGAGAAATTGATTATGATGAGAGTGCACAACACTAATTAATGAGCCAGCT 480
    |||||||
DB 435 AANAACATATTGAGAAATTGATTATGATGAGAGTGCACAACACTAATTAATGAGCCAGCT 494
QY 481 GTTGGCATGTGTTTCTTACTTAGTCTCCCAA-GGAAAACTCTAACTGAATCTTCAG 539
    |||||||
DB 495 GTTGGCATGTGTTTCTTACTTAGTCTCCCAAAGGAAAACTCTAAATTTGAATCTTCAG 554
QY 540 CNGAATAA-CCTTAATATATCTTTGTTAGCCCAACAAAAA-CCTTTTGTTTACATAGTTC 597
    |||||||
DB 555 CAGAAATAATCCTTAATATATCTTTGTTAGCAAAAAAAGCCTTTTGTTTACATAGTTC 614
QY 598 TTT-GGATTTTACTGTCTCTAATTTTATTTCTGAAACTCCATTTTCCCGAGACCATTAAT 656
    |||||||
DB 615 TTTGGGATTTTACTGTCTCTAATTTTATTTCTGAAACTCAATTTTAACTCCAGACCATTAAT 674
QY 657 ACCCTAATTAACCTTTGTTATGACAGTGT 687
    |||||||
DB 675 ACCATA-TTAACTTTGTTNTGACAGTGT 704
```

```

RESULT 4
US-09-796-692-7228
; Sequence 7228, Application US/09796692
; Publication No. US20020198362A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7228
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7228
```

```

Query Match      68.6%; Score 471; DB 10; Length 506;
Best Local Similarity 99.4%; Pred. No. 8.8e-76;
Matches 471; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```

QY 1 ATTTAATTTTCTATTAACAATCTCTCAAGCATTATTTATCCTATATCTCACTGAA 60
    |||||||
DB 33 ATTTAATTTTCTATTAACAATCTCTCAAGCATTATTTATCCTATATCTCACTGAA 92
QY 61 TTTTAAATAATAACATTAGTATTTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 120
    |||||||
DB 93 TTTTAAATAATAACATTAGTATTTAGAAAACTAGAAAAAAGATNAATGCAGATAATTAA 152
QY 121 ACTTACATGAAAAAGAAAAATTATTAACAAGAGCTGAGAACGTTATAATTGAATGAGA 180
    |||||||
DB 153 ACTTACATGAAAAAGAAAAATTATTAACAAGAGCTGAGAACGTTATAATTGAATGAGA 212
QY 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 240
    |||||||
DB 213 TTATAATTGAAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATCTTAATGATGG 272
QY 241 TGTTTATGACTAATACATGATTTTTCATAAGGAACCCAGTTAAAAATATTTTAT 300
    |||||||
DB 273 TGTTTATGACTAATACATGATTTTTCATAAGGAACCCAGTTAAAAATATTTTAT 332
QY 301 TTTTAAATAAAGCCTGTGTTCAGACCTCTGATCATATTTCTTTAATTTTGATTTGGGAAGA 360
    |||||||
DB 333 TTTTAAATAAAGCCTGTGTTCAGACCTCTGATCATATTTCTTTAATTTTGATTTGGGAAGA 392
QY 361 AAATACGTCTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 420
    |||||||
DB 393 AAATACGTCTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCNCCTAATTTT 452
```


DB 433 AAGACTATTGAGAAATGATTATGACATGCAAGTGCAACACTAATTACTGG 486

RESULT 7
US-09-835-992A-22

; Sequence 22, Application US/09835992A
; Patent No. US20020037541A1
; GENERAL INFORMATION:
; APPLICANT: Obata, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AN
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING GASTRIC CANCER
; FILE REFERENCE: L0461/7112
; CURRENT APPLICATION NUMBER: US/09/835,992A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 08/896,164
; PRIOR FILING DATE: 1997-07-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (12)..(12)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (19)..(19)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (33)..(33)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (35)..(35)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (42)..(42)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (47)..(47)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (56)..(56)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (64)..(65)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (69)..(69)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (71)..(71)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (74)..(74)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (78)..(78)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (80)..(80)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (84)..(84)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (92)..(93)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (95)..(95)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure

; LOCATION: (107)..(107)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (122)..(122)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (130)..(130)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (139)..(139)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (148)..(148)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (157)..(157)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (164)..(164)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (169)..(169)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (174)..(174)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (176)..(176)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (179)..(179)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (185)..(185)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (189)..(189)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (191)..(191)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (200)..(200)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (203)..(203)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (208)..(208)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (213)..(214)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (216)..(216)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (225)..(225)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (229)..(229)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (246)..(246)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (249)..(249)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (255)..(256)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (260)..(260)

```

; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (262)..(263)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (270)..(270)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (274)..(274)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (287)..(287)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (292)..(292)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (296)..(298)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (305)..(305)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (312)..(312)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (329)..(329)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (332)..(332)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (342)..(342)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (348)..(348)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (357)..(357)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (379)..(379)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (381)..(381)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (387)..(388)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (394)..(395)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (399)..(400)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (406)..(406)

```

Query Match 59.9%; Score 411.8; DB 9; Length 843;
Best Local Similarity 71.8%; Pred. No. 4.9e-65;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

```

QY 1 ATTTAATTTCTAATAACATTTCTCAAGCATTTATCTCACTGAA 60
    |||||
Db 15 ATTTAATTTCTAATAANCNTCTCCNCAANCAATTATTNACCTATTNNCNCNGAN 74
    |||||
QY 61 TTTAANAATAACATTAAGTATTGAAAAAAGATNAATGAGATTAATTAA 120
    |||||
Db 75 TTTNANAANTACTTTNNNTTAAAAAACCTNGAAAAAATAATNGCAATANTTAA 134
    |||||

```

```

QY 121 ACTTACATGAAAAAGAAATTTATACAAAGACTGAGACGTTATAAATGAAATGAGA 180
    |||||
Db 135 CCTTNCCTGAAAAANGAAATTTNTACCAANGGACNGAAANCNTNTAATTNGAANTNAAA 194
    |||||
QY 181 TTATAATTTGAAAACCTGCATCTGAAAGCAACTTTATTGTTCAATTAATTAATGATGG 240
    |||||
Db 195 TTATANTTNGAAANGCGCNCNNGAAGAACCAANCTTNTATGTCATTAATCTCTTNAANGAGGG 254
    |||||
QY 241 TGTTTTACTATACACTGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT 300
    |||||
Db 255 NNTTANNACTAATNCCNGAATTTCCAAATANGAANCCNNNTTAAANTNTTTTNAAT 314
    |||||
QY 301 TTTAAAAATAGCCTGTGTTCAGCTCTGATCATATTTCTTTATTTGATTGGGAAGA 360
    |||||
Db 315 TTTAAAAATACCCNGTNTCCAAACCCCGAATCANATTCCTTNTATTTGATTGGGAAAAA 374
    |||||
QY 361 AAATACTGTTTGTGATGACATGAATGCAAAATTTTAGATTTTAAATCTNCTAATTTT 420
    |||||
Db 375 AAATNCNGTCCNNATACCNNGAANNCAANTTTTAAATTTTAAACCCCTANTTTT 434
    |||||
QY 421 AAGACTATTTGAGAAATTTGATTAATGACATGAAGTGCACACACTAATTAATGAGCCAGCT 480
    |||||
Db 435 AAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCTANTTNCNGCCACCN 493
    |||||
QY 481 GTTGCAATGTGTTTCTTACTAGTTCCTCCGAAGAAAACCTTTAACTGAATCTTGACG 540
    |||||
Db 494 GTGGCNTNGTNTCTTACTTANTCCCCCAAGAAANNCTTAAANCANCTCCNCC 553
    |||||
QY 541 NGATTAACCTTAATATACTTTGTTAGCCAAAC--AAAACTTTTGTGTTTACATAGTTCT 598
    |||||
Db 554 AAAATAACCTTAANTATCTTGTTAGCAAAANCAAAACGTTTNGTTTACNTANTCCT 613
    |||||
QY 599 TTGATTTTACTGTTCTTAATTTTATTTCTGAAACTCCATTTTCCCGACCATTAATTAC 658
    |||||
Db 614 TGGATTTAACGGGTCCCAATTTNATCCNGAACCCANTTTTCCCGCAACATANTTAC 673
    |||||
QY 659 CCTATTAACTTTGTTATGCACAGTTGTT 687
    |||||
Db 674 CAT-TTTACCTTGTAAGCNCAGTNGTT 701
    |||||

```

RESULT 8
US-10-102-524-161/c
; Sequence 161, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-161

Query Match 23.6%; Score 161.8; DB 14; Length 510;
Best Local Similarity 94.8%; Pred. No. 5e-20;
Matches 199; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

```

QY 479 CTGTTGCAATGTGTTTCTTACTTACTTCCCAAGAAAACTTTAACTGAATCTTCA 538
    |||||
Db 510 CTGTTGCAATGTGTTTCTTACTTACTTCCCAAGAAAACTTTAAATGAATCTTCA 451
    |||||
QY 539 GCNGAATAA-CCTTAATATACTTGTGTAAGCAACAAAA-CITTTTGTGTACATAGTT 596
    |||||

```



```
RESULT 12
US-10-240-485-151
; Sequence 151, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 151
; LENGTH: 7676
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-151
```

```
Query Match          9.7%; Score 66.8; DB 12; Length 7676;
Best Local Similarity 48.5%; Pred. No. 0.016;
Matches 210; Conservative 0; Mismatches 220; Indels 3; Gaps 1;

QY 21 ATTCTCTCAAGCAATTTTATTCCTATATCTCAGTAATTTTAAANAATACATT-- 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3697 ATTTGTTAATATAATATATATATTAGTGTAGTAAATTTAAATGTTAATAATTAA 3756

QY 78 AGTATTAGAAAACTAGGAAAAAGATNAATGCAGTAATTAACTTACATGAAAAAGGA 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3757 TATATTAAATTATATATTAATAATATAAGATATGAAATATTTAAGAAATGAAAACTT 3816

QY 138 AAATTATACAAAGCAGTGAACGTTATAAATGAAATGAGATTATTAATTTGAAAACTG 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3817 AATGTATGTTAAATTTTATTATATAGGAATATTTTATTATAGATATAGAAAAATT 3876

QY 198 CATCTGAAGCAAACTTATGTTCAATATTCCTAATGATGCTTTTATGACTAATAC 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3877 AATTTAAATTAATAAATTTAGTTTTTTTATGATAAAGTTTTTTTATTAATTAATAA 3936

QY 258 ACTGATTTTCAATAGGAAGAACCATGTAAATAATATTTTATTTAATAAAGCCTGT 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3937 GATAATTTTTTTTTTTTAAAGTAGAGTTTATTAATTTTAGTGAGTTATTAAGA 3996

QY 318 GTTCAAGCTCTGATCATATTTCTTATTGATTTGGGAAGAAATACGTCTTGATA 377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3997 TTTTGGTGIGTTTTTTTTTATTGAAATATTTTTGTTGAAATTAATTTAGTTTA 4056

QY 378 GCATGAATGCAAAATTTTAGATTTTAAATCNCNTAATTTTAAAGACTATGAGAAAT 437
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4057 AAATAGTAATTAATTAATTTTATTATTATTTTATAGTAATAGTAATAATTTAGAGATT 4116

QY 438 TGATTAATGACAT 450
   ||| ||| |||
Db 4117 AGTAGAATAGTAT 4129
```

```
RESULT 13
US-10-311-455-2147/c
; Sequence 2147, Application US/10311455
```

```
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2147
; LENGTH: 113515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2147
```

```
Query Match          9.5%; Score 65; DB 12; Length 113515;
Best Local Similarity 44.8%; Pred. No. 0.083;
Matches 282; Conservative 0; Mismatches 344; Indels 3; Gaps 1;

QY 5 TAAATTTCTATTAACATTTCTCTCAAGCATTAATTTTATCCTATATCTCAGTAATTTT 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45803 TAAAAATATACCAAAAATTTTATTTAACTCAAAAAATAATCATATATATACATTAATTCA 45744

QY 65 AANAATAACATTAAGTATTAAGAAAACTAGGAAAAAGATNAATGCAGTAATTAACCT 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45743 AAAAAAAATCCCATTAATTAATTAACATTAACATTAACATTAATTAATTAACATTA 45684

QY 125 ACATGAAAAAGAAAAATTATTAACAAGACTGAGAACGTTATTAATGAAATGAGATTAT 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45683 AATAAATAAATAAATAAATAAATAAATGAATTAATTAATTAATTAATTAATTAATTA 45624

QY 185 AATTGAAAACTGACTGGAAGCAAACTTATGTTCAATTAATGTAATGATGCTT 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45623 AATACTAATACTAAATAAATTTTACAATAATATACAAATATTAATTTCTTAATTA 45564

QY 245 TTATGACTAATACACTGATTTTTCATTAAGGAAACCATGTTAAATAATTTTATTTTA 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45563 ATCAATACATTC---CACTAACAACAATAATATTACATTTTCATTAACAACAATAATTT 45507

QY 305 AAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTATTGATTTGGGAAGAAAT 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45506 TCATTATCAATTTTCTACCATTTACAATAAATAAATAATTTCAAAACACTAA 45447

QY 365 ACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAAATCTCNCNTAATTTTAAGA 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45446 ACAAAACTTAATAAATAATACATACAAAAAATAATCTTTACTTTTACCAATTTTCAAA 45387

QY 425 ACTATTGAGAAATGATTAATGACATGAGAGTGACACAACACTAATTAATCTGAGCAGCTGTTG 484
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45386 TCTTTACAAATTTTCTTAACCTTAATCATCACATTTTAAACACTAAATTTTCAAAAAATTA 45327

QY 485 GCATGTGTTTCTTACTTAGTCTCCCAAGAAAACTCTTAAAGTGAATCTTCAGCNGAA 544
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45326 ACTATAATTAATTTAAATTTTTCCTTCACTATTAATTTACTATTAATTTCAAAATCCC 45267

QY 545 TAACCTTAATAATACTTTGTTAGCAACAACAATTTTGTGTTACATAAGTCTTTGGAT 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45266 TACAATTTAACAACAATATTTCTATTAACATATATCAAAAACTATTCATTTTATTTAACT 45207

QY 605 TTTACTGTTCTTAATTTTATTTGAACT 633
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 45206 TTAATATATTTAAAAATATATCTTTATATT 45178
```


Search completed: November 27, 2003, 12:52:57
Job time : 263.057 secs

```
|||||
Db 61 TTTTANAATAACATAGTATTAGAAAACTAGGAAAAAGATMAATGCAGATTAATTA 120
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATTAATTGAAATGAGA 180
Db 121 ACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATTAATTGAAATGAGA 180
QY 181 TTATATTTGAAAACTGATCTGAAAGCAAACTTTATGTTCAATTATTTCTTAATGATGG 240
Db 181 TTATATTTGAAAACTGATCTGAAAGCAAACTTTATGTTCAATTATTTCTTAATGATGG 240
QY 241 TGTATTATGACTAATACATGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT 300
Db 241 TGTATTATGACTAATACATGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT 300
QY 301 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTGGGAAGA 360
Db 301 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTGGGAAGA 360
QY 361 AAATACGTGTTCTGATGACATGAAATGCAGAAATTTTATGATTTTATCTCNCCTAATTT 420
Db 361 AAATACGTGTTCTGATGACATGAAATGCAGAAATTTTATGATTTTATCTCNCCTAATTT 420
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTACTGGCCAGCT 480
Db 421 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTACTGGCCAGCT 480
QY 481 GTTGCATGTGTTCTCTACTAGTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC 540
Db 481 GTTGCATGTGTTCTCTACTAGTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGC 540
QY 541 NGATAACCTTAATATACTTTGTTAGCCAAACAACCTTTTGTTTACATAGTTCTTT 600
Db 541 NGATAACCTTAATATACTTTGTTAGCCAAACAACCTTTTGTTTACATAGTTCTTT 600
QY 601 GGATTTTACTGTTCCCTAATTTTATTTCTGAAACCTCAATTTTCCCGAGACCATTAATACC 660
Db 601 GGATTTTACTGTTCCCTAATTTTATTTCTGAAACCTCAATTTTCCCGAGACCATTAATACC 660
QY 661 TATTTAACCTTGTATGACAGTTGTT 687
Db 661 TATTTAACCTTGTATGACAGTTGTT 687

RESULT 2
US-08-896-164-19
; Sequence 19, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
```

```
REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-896-164-19

Query Match 89.1%; Score 611.8; DB 3; Length 714;
Best Local Similarity 97.1%; Pred. No. 3.4e-115;
Matches 671; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

QY 1 ATTTAATTTTCTATTAAACATTTCTCAAAAGCATTATTTTATCTATATCTCACTGAA 60
Db 15 ATTTAATTTTCTATTAAACATTTCTCAAAAGCATTATTTTATCTATATCTCACTGAA 74
QY 61 TTTAANAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGAGATAATTAA 120
Db 75 TTTAAGAATAACATTAGTATTAGAAAACTAGAAAAAAGATNAATGAGATAATTAA 134
QY 121 ACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATTAATTGAAATGAGA 180
Db 135 ACTTACATGAAAAAGAAAAATTATACAAAGAGCTGAGAACGTTATTAATTGAAATGAGA 194
QY 181 TTATATTTGAAAACTGATCTGAAAGCAAACTTTATGTTCAATTATTTCTTAATGATGG 240
Db 195 TTATATTTGAAAACTGATCTGAAAGCAAACTTTATGTTCAATTATTTCTTAATGATGG 254
QY 241 TGTATTATGACTAATACATGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT 300
Db 255 TGTATTATGACTAATACATGATTTTTCATTAAGGAAACCCATGTTAAAAATATTTTAT 314
QY 301 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTGGGAAGA 360
Db 315 TTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTGGGAAGA 374
QY 361 AAATACGTGTTCTGATGACATGAATGCAAAATTTTATGATTTTATCTCNCCTAATTT 420
Db 375 AAATACGTGTTCTGATGACATGAATGCAAAATTTTATGATTTTATCTCNCCTAATTT 434
QY 421 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTACTGGCCAGCT 480
Db 435 AAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTACTGGCCAGCT 494
QY 481 GTTGCATGTGTTCTTACTAGTCTCCCAA-GGAAAACTCTTAACCTGAATCTTCAG 539
Db 495 GTTGCATGTGTTCTTACTAGTCTCCCAAAGGAAAACTCTTAATGAACTTCAG 554
QY 540 CNGAATAA-CCTTAATATACCTTTGTAGCCAAACAACAA-CCTTTTGTGTTACATAGTTC 597
Db 555 CAGAATAATCCTTAATATACCTTTGTAGCCAAACAACAAAGCTTTTGTGTTACATAGTTC 614
QY 598 TTT-GGATTTTACTGTTCTTAATTTATTTCTGAAACTCCATTTTCCCGAGACCATATT 656
Db 615 TTTGGGATTTTACTGTTCTTAATTTATTTCTGAAACTCAATTTTACCCGAGACCATATT 674
QY 657 ACCCTATTTAATTTGTTATGACAGTTGTT 687
Db 675 ACCATA-TTAACTTTGTTNTGACAGTTGTT 704

RESULT 3
US-08-896-164-22
; Sequence 22, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
```

TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,164
FILING DATE: July 17, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6218521man D. Hanson
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 843 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-896-164-22

Query Match 59.9%; Score 411.8; DB 3; Length 843;
Best Local Similarity 71.8%; Pred. No. 6.8e-75;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

QY 1 ATTTAATTTCTATTAAACATCTTCTCAAGCATTTATCTTATCTCACTGAA 60
DB 15 ATTNAATTCCTATTAAACNTCTCCNCAANCAATTATTNACCTTNNCNCNGAN 74

QY 61 TTTTAAANAATAACATTAGTATTAGAAAACTAGAAAAAGATNATGAGATAATTAA 120
DB 75 TTTNANAANAATACCTTTNTNTTAAAAAACCINGAAAAAATAATNGCAATANTTAA 134

QY 121 ACTTACATGAAAAAGAAAATTATACAAAGACTGAGACGTATATAATTGAATGAGA 180
DB 135 CCTTNCCTGAAAAANGAATTTNTACCAANGACNGAANCNTTNTAATTNGAANTMAA 194

QY 181 TTATAATTGAAAACTGCATCTGAAAGCAACTTATTGTTCAATTATCTTAATGATGG 240
DB 195 TTATANTTNGAAACGGCNCNGAACAACCAACTTNATGTCCAATTATCTTNAANGAGGG 254

QY 241 TGTTTATGACTAATAACAATGATTTTCAATAAGAAAAACCATGTTAAAAATATTTTAT 300
DB 255 NNTTANNACTAATNCCNGATTTTCCATANGAANCCNNNTTAAANTNTTTTAT 314

QY 301 TTTAAAAATAAGCCTGTGTCAAGCTGTGATCATATTTCTTTATTTGATTTGGGAGA 360
DB 315 TTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCCTTNTATTTGATTTGGGAAAA 374

QY 361 AATACTGTTTCTGATAGCATGAATGCAAAATTTTGAATTTTAAATCTCNCATAATTTT 420
DB 375 AATTCNGTTCNNATACCCNNGAANNCAAAATTTTAAATTTTAAACCCCTANTTTT 434

QY 421 AAGAACTATGGAATAATGATTAATGACATGAAGTGACACACATAATTAATGCGCAGCT 480
DB 435 AAAAANTATNGAAAAANTNGATTANNGACTTGAAATTGC-CAACCTANTTNCNGGCCACCN 493

QY 481 GTTGGCATGTGTTTCTTACTAGTTCTCCCAAGAAAACTTTAAACTGAATCTTCAGC 540
DB 494 GTGGCNTNGTNTCTTACTTANTCCCCCAAGAAAAANNCTTAANCNGAANTCTCNCNC 553

QY 541 NGAAATACCTTAATATATACCTTGTAGCCAAAC--AAAACCTTTTGTGTACATAGTCT 598
DB 554 AAAATACCCCTTAANTATCCTGTGTAACCAANCAAAAACCTTTTNGTTACNTANTCCT 613

QY 599 TTGATTTTACTGTTCTCTAATTTTATTTCTGAACACTCCATTTTCCCGACACATATTAAC 658
DB 614 TGGGATTTAACGGGTCCCAATTTNATCCNGAACCCANTTTTCCCCNAACCATANTTAC 673

QY 659 CCTATTTAATTTGTTATGACACAGTTGT 687
DB 674 CAT-TTACCTTGTAAGCNCAGTNGTT 701

RESULT 4
US-09-313-294A-1553/c
Sequence 1553, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Ialguudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
NUMBER OF SEQ ID NOS: 1999-05-14
SOFTWARE: PERL Program
SEQ ID NO 1553
LENGTH: 173
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700551132H1
US-09-313-294A-1553.

Query Match 23.4%; Score 161; DB 4; Length 173;
Best Local Similarity 98.8%; Pred. No. 2e-24;
Matches 161; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TTTAATTTCTATTAAACATCTTCTCAAGCATTTATCTTATCTCACTGAAAT 62
DB 163 TTTAATTTCTATTAAACATCTTCTCAAGCATTTATCTTATCTCACTGAAAT 104

QY 63 TTAANAATAACATTAGTATTAGAAAACTAGAAAAAGATNATGAGATAATTAAAC 122
DB 103 TTAAGAAATACATTAGTATTAGAAAACTAGAAAAAGATNATGAGATAATTAAAC 44

QY 123 TTACATGAAAAAGAAAATTATACAAAGACTGAGAACGTTA 165
DB 43 TTACATGAAAAAGAAAATTATACAAAGACTGAGAACGTTA 1

RESULT 5
US-07-867-106-2/c
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; FILING DATE: 02-NOV-1989
; APPLICATION NUMBER: PCT/AU90/00530
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
;
; NAME/KEY: CDS
; LOCATION: 2378..5038
;
; US-07-867-106-2

Query Match          9.7%; Score 66.4; DB 1; Length 5852;
Best Local Similarity 46.8%; Pred. No. 3.2e-05;
Matches 240; Conservative 0; Mismatches 269; Indels 4; Gaps 1;

QY 2 TTTTAATTTCTAATTAACATCTCTCAAGCATTTATTTATCTCAATCTCACTGAAT 61
   |||||
DB 5751 TTGTATTTTATATATATGTTATTTGTTGTTGTTTCTTACTTAATTTCTAATTTT 5692

QY 62 TTTAANAATTAACATTAAGTATTAGAAAACTAGAAAAAGATNAATGAGATAATTAA 121
   |||||
DB 5691 TTTATAAATTAATTAATTAATTTAATAATAAAAAAATAAAAAAATAAAAAA 5632

QY 122 CTTACATGAAAAAGAAATTTATACAAAGAGCTGAGAACGTTTAAATGAATGAGAT 181
   |||||
DB 5631 ATTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAA 5572

QY 182 TATAATTTGAAAACTGATCTGAAGCAACTTATTTGTTCAATTAATCTTAATGATGT 241
   |||||
DB 5571 TAAAAAATGTAAGGGTTTTTTTAAATTAATGATCGATTTTATTTAAATCATTTGA 5512

QY 242 GTTTATGACTAATACACTGATTTTCAATAAGAAACCATGTTAAATAATTTTATTT 301
   |||||
DB 5511 CGAGATTAAAAAATCTTAACAATTAACATATTGATTTTCTTTTCTTTTCTTTT 5452

QY 302 TT---AAAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATTTTGATTTGGA 357
   |||||
DB 5451 TTTTAAAAATTCAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTTTA 5392

QY 358 AGAAATATCTGTTCTGATAGCATGAAGAAATGAAATTTTAGATTTTAAATCTCNCCTAAT 417
   |||||
DB 5391 TAAATTTTAATTAATTAATTTTAAAAATTAATGATCTATGATCTAATAATTTTCCATG 5332

QY 418 TTTAAGAACTATTGAGAAATTTGATTAATGACATGAGTGCACACACTAATTAATGAGCA 477
   |||||
DB 5331 TTTTAAAAATTTTAAAGTATTTTAAATTAATTAATTAATTAATTAATTAATTAATGGA 5272

QY 478 GCTGTGGCATTTGTTCTTCTTACTAGTTCTCC 510
   |||||
DB 5271 AATGATACTAGATCTTTTCCCAAGTTTCC 5239
```

```

RESULT 6
US-09-641-638-651/c
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; NAME/KEY: exon
```


OY	189	TGAAACTGCATCGAAGCAAACTTTATGTTCATTATTCCTTAATGATGGTGT	248
Db	11359	AAAAATTAAAAATAATTAATAATTTAATTTAATAATTAATAATTAAGTTAAATTAAT	11300
OY	249	GACATAACACCTGATTTTTCATAAGGAACCACATGTTAAAAATATTTTAT--TTTAAA	306
Db	11299	ATTAATTTAAATTAAGTTAAATTAATATTTAAATTAATTTAATTTAATAATTAATAA	11240
OY	307	AATAAGCCTGTGTTCAGCTCTGATCATATTTCTTTATTTTGATTTGGAGAATAATAC	366
Db	11239	ATTTAAATGTTATTTTAAAAATTAAATTTAAATTTTAAATTTAAATTTTAA	11180
OY	367	TGTTTCGATAGCATGAATGCAAAATTTTAGATTTTAAATCTCNCTAATTTTAAGAAC	426
Db	11179	TTATTTATTTTAAATTTAAAAATTTTAATTAATATTTAAATTTAAATTTTAATTTAATTT	11120
OY	427	TATTGAGAAATGATTTAATGACATGAAGTCACAACACTAATTTACTGCCACGTGTTGCC	486
Db	11119	AAATTAATAATTTAATAATTTAAAAATATTAATAATATTTCCCTTAGCTATTAAAAAAGGC	11060
OY	487	ATTGTGTTT	495
Db	11059	ATGTCCTT	11051

```

RESULT 7
US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
;
; GENERAL INFORMATION:
;
; APPLICANT:  Philippsen, Peter
; APPLICANT:  Pohlmann, Rainer
; APPLICANT:  Steiner, Sabine
; APPLICANT:  Mohr, Christine
; APPLICANT:  Wendland, Jurgen
; APPLICANT:  Knechtle, Philipp
; APPLICANT:  Rebischung, Corinne
;
; TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;
; TITLE OF INVENTION:  AND USES THEREOF
;
; NUMBER OF SEQUENCES:  1152
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE:  No. 6239264artis Corporation
;
; STREET:  3054 Cornwallis Road
; CITY:  Research Triangle Park
; STATE:  No. 6239264th Carolina
;
; COUNTRY:  USA
;
; ZIP:  27709
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE:  Floppy disk
;
; COMPUTER:  IBM PC compatible
;
; OPERATING SYSTEM:  PC-DOS/MS-DOS
;
; SOFTWARE:  PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER:  US/08/998,416
;
; FILING DATE:  24-DEC-1997
;
; CLASSIFICATION:  435
;
; PRIORITY APPLICATION DATA:
;
; APPLICATION NUMBER:  CH 0016/97
;
; FILING DATE:  31-DEC-1996
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME:  Meigs, J. Timothy
;
; REGISTRATION NUMBER:  38,241
;
; REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE:  919-541-8587
;
; TELEFAX:  919-541-8689
;
; INFORMATION FOR SEQ ID NO:  288:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH:  837 base pairs
;
; TYPE:  nucleic acid
;
; STRANDEDNESS:  single
;
; TOPOLOGY:  linear
;

```

```

; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1241RP
US-08-998-416-288

```

Query Match	7.8%	Score 53.6;	DB 3;	Length 837;
Best Local Similarity	45.6%;	Pred. No. 0.01;		
Matches 260; Conservative	0;	Mismatches 303;	Indels 7;	Gaps 2;

[illegible]

```

RESULT 8
US-09-134-001C-150/c
; Sequence 150, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 150
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-150

```


TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
; OTHER INFORMATION: plasmid RP4"
US-09-056-075-1

Query Match 7.5%; Score 51.4; DB 2; Length 6243;
Best Local Similarity 46.8%; Pred. No. 0.034;
Matches 191; Conservative 0; Mismatches 214; Indels 3; Gaps 1;

QY 65 AANAATACATAGTATTAGAAAACTAGGAAAAAGATNAATGCAGATTAATTAACCTT 124
Db 1234 AAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAAT 1293
QY 125 ACATGAAAAAGAAATTAATACAAAGGACTGAAACGTTATTAATTGAAATGAGATTAT 184
Db 1294 AAAAAAATATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAA 1353
QY 185 AATTGAAACTGCATCTGAAAGCAAACTTTATTGTTCAATTATTCTTAATGATGTTGT 244
Db 1354 AAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAAT 1413
QY 245 TTATGACTAATACATGATTTTTCATAAGAAACCCATGTTAAATATTTTATTTTA 304
Db 1414 TTTAAAGTTGAAAAAAATTTTATATATTAATCTTTGAGAAAAAGA--ATATAA 1470
QY 305 AAAATAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATTGATTGGGAGAAAAAT 364
Db 1471 AAAATGAGCCTTAATAAAGCCCATTTTTCATATACGTAATATGACGTTCTTAATGTT 1530
QY 365 ACTGTTCTGATAGCATGAAATGCAAAATTTTAGATTTTTAATCTCNCNTAATTTTAA 424
Db 1531 TTTATTGCTACTCTTAACATTAGAGTAATTTCTTAATTTTAAAGCCTTTTCTTTAAG 1590
QY 425 ACTATTGAGAATGATTAATGACATGAAGTGCAACAACCTAATTACT 472
Db 1591 GCTTTATTTTTCTTAATACATTAACTCTTTTGTGTTGCT 1638

RESULT 11

US-08-998-416-186/c

; Sequence 186, Application US/08998416

; Patent No. 6239264

; GENERAL INFORMATION:

; APPLICANT: Philippesen, Peter

; APPLICANT: Pohlmann, Rainer

; APPLICANT: Steiner, Sabine

; APPLICANT: Mohr, Christine

; APPLICANT: Wendland, Jurgen

; APPLICANT: Knechtel, Philipp

; APPLICANT: Reibischung, Corinne

; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 1152

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6239264artis Corporation

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: No. 6239264th Carolina

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074RP
US-08-998-416-186

Query Match 7.5%; Score 51.2; DB 3; Length 615;
Best Local Similarity 46.5%; Pred. No. 0.03;
Matches 265; Conservative 0; Mismatches 297; Indels 8; Gaps 3;

QY 6 AATTTCCTAATTAACATCTCTCTCAAGCATTTTATCTCTATATCTCACTGAATTTTA 65
Db 573 AATATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 514
QY 66 ANAATAACATTAATTAAGAAAACTAGGAAAAAGATNAATGCAGATTAATTAACCTTA 125
Db 513 TTAATTAATTAATAAAATTAATTAATTAAGAAATTAAGCTTAAATTAATTAATTAAT 454
QY 126 CATGAAAAAGAAAAATTAATACAAAGAGCTGAGAACGTTATTAATGAATGAGATTATA 185
Db 453 CTTATTAAGAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 394
QY 186 ATTGAAACTGC--ATCTGAAGCAAACTTTATTGTTCAATTATCTTAATGATGTG 242
Db 393 AAAATAATTAATTAACATATTTAATAATAATTAATCTTTATTAATAATAATAATATAT 334
QY 243 TTTATGACTAATAC--ACTGATTTTTCATAAGAAACCCATGTTAAAAATATTTTAT 300
Db 333 TTAATAACAATTAATAATAATAATAATAATTAATTTGATTAATCTAATTAATTTATTA 274
QY 301 TTTAAATAAAGCCTGTGTTCAAGCTCTGATCATATTCTTTTATTGATTGGGAGAGA 360
Db 273 AGAAATTAATATATCTAATAATAATTTTAATACTAATTTAAATTTGAACATAGACTAA 214
QY 361 AAATACTGTTCTGATAGCAATGCAAAATTTTGAATTTTAAATCTCNCNTAATTTT 420
Db 213 ATAGTATCATATTAAATATATTTTATTAATTAATTAATTAATTAATTAATTAATTA 154
QY 421 AAGACTATTGAAATTTGATTAATGACATGAAGTGCAACAACCT--AATTACTGGCCA 477
Db 153 GTAATTAATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 94
QY 478 GCTGTGCAATGTTGTTCTTAAGTTCTCCCAAGAAACCTCTTAACCTGAATCTTC 537
Db 93 ATGATATAATAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 34
QY 538 AGCNGAATAACCTTAATATATCTTTGTTAG 567
Db 33 ATCTTTAATAATTAATTAATTAATTAATTTAG 4

RESULT 12

```

US-09-601-198-63
; Sequence 63, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Elison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Helner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-63

Query Match          7.4%; Score 51; DB 4; Length 1692;
Best Local Similarity 47.0%; Pred. No. 0.036;
Matches 156; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY      118 TAACTTACATGAAAAAGAAAATTATACAAAGCACTGAGAACGTATTAATGGAATG 177
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      189 TACAGTAAACAAAAATTAACAATTATTAATAAATACCTAATTGCAATTAAATGAATT 248

QY      178 AGATTATTAATTGAAAACCTGCATCTGAAGCAAACTTATGTTCAATTATCTTAATGA 237
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      249 AGATTTTCATAAAAAAGCTAATTAATTACAAAGATTATCTGAAAATCATAAATTATT 308

QY      238 TGGTGTATTATGACTAATACACTGATTTTTCATAAGAAACCATGTTAAAAATATTTT 297
      ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      309 CTCTTATTATCACTTATTAAGTTTATTTAAGCAAAATGAATTAATAATTAATAATATGT 368

QY      298 TATTTAAAAATAAGCCGTGTCAGCTCTGATCATATTCTTTATTTGATTGGGA 357
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      369 AGATGAATTGTAGATTAAATCAATCTTTTGAATGATTTTATTATTAATGATCAACA 428

QY      358 AGAAATACTGTTTCTGATAGCATGAATGCAAAATTTTGAATTTTAAATCNCNTAAT 417
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      429 AGAAAGAAAAATTGAAGATVACCATAGAATAATTAGATTAAATGATTAAGATCTTAAG 488

QY      418 TTTAGAACTATTGAGAAATTGATTATGACA 449
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      489 TAAATATAGTATTCCAAATTTAAATTATCAAA 520

RESULT 13
US-08-213-419B-3/c
; Sequence 3, Application US/08213419B
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Inselburg, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: J11-002CNCIP
; CURRENT APPLICATION NUMBER: US/08/213,419B
; CURRENT FILING DATE: 1994-03-14
; PRIOR APPLICATION NUMBER: US 07/870,506
; PRIOR FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407) .. (2439)
; NAME/KEY: CDS
; LOCATION: (2598) .. (3404)
; NAME/KEY: CDS
; LOCATION: (3580) .. (3720)
; NAME/KEY: CDS
; LOCATION: (3850) .. (5835)
US-08-213-419B-3

Query Match      7.3%; Score 50.2; DB 4; Length 6124;
Best Local Similarity 48.9%; Pred. No. 0.06;
Matches 133; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY    77 TAGTATTAGAAAAAAGGAAAAAGATNAAATGCAGATAATTAACCTTACATGAAGAAG 136
     ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    6077 TATTTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAGATTATATTACATGAATAAA 6018

QY    137 AAAATTATACAAGAGACTGAGAACGTTATAATTGAATGAGATTATATTGAAAAC 196
     ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    6017 ATAAATTCTTAAGAAAAAATAATATAATATAATTTATATATATTATAGAGGTACAT 5958

QY    197 GCATCTGAAGCAAACTTATTGTTCAATTATTCCTTAATGATGCGTTTATGACTAATA 256
     ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    5957 GATAAATTAAAGATATTATCTACCCTTAATAAAGAAATGAGAGATTATGCCCTAAT 5898

QY    257 CACTGATTTTTCAATAAGAAACCAGTTAAATAATTTTATTTTAAAAATAAGCCTG 316
     ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    5897 TAAAAATTAAAGCATACATAAAAAAATAATTTTATTTTGGTTATATTAATT 5838

QY    317 TGTTCAAGCTCTGATCATATTCTTTTATTTT 348
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    5837 TATACATAACAGAAATACATTGCTTATTGCT 5806

RESULT 14
US-08-998-416-1137/c
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPIT
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241

```

```

; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAGI692RP
;
US-08-998-416-1137

```

Query Match	7.2%;	Score 49.2;	DB 3;	Length 636;
Best Local Similarity	45.8%;	Pred. No. 0.076;		
Matches 240;	Conservative 0;	Mismatches 277;	Indels 7;	Gaps 2;

QY	47	TATATCTCACTGAATTTTAA	AAAAATAACA	CTAGTATTAG	AAAAAACTAG	AAAAAAGATNA	106
Db	523	TATTAATTTATTAATAATAT	AAAAAATTAA	TAGA	AAATTTAAAGTTAA	AAATTTTAA	464
QY	107	ATGCAGATTAATTAACCTT	TACATGAAAAAG	AAATTTAACA	CAAGCACTGACA	AGCTTAT	166
Db	463	ATAATAATTCCTTATTA	AAAAAGATTA	AAATATA	TAAATCAACAT	ATATTTTATA	404
QY	167	AAATTGAATGAGATTAT	TAATTTGAAA	ACTGCATCTG	AAAGCAA	CTTTATGTTCAATT	226
Db	403	TATTATAATAAAAAAT	ATAATATTAC	CAATATTT	TAAATTA	AAATCTTTATATA	344

```

QY      227  ATTCTTAATGATGGTGTTTATGACATAAACAAGTATTTTCAATTAAGGAACCCATGTT 286
          ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      343  ATATATTATTATTAATAACAATTAAATAA-----TAATATATTAAATTGATTAATCTATT 288

```

287 TAATAATTATTAAAGAAATAATAATCTAATAATTTTAATACTAATTAAATT 228

Db 227 TGACATAGACTAAATAGATTCATATTAAATATTATTTTATTAATTAATAAATTAA 168

Db 167 TAATGATGAATTAACTAATTAATAATTATATATAAGTATTAAATCAATAAT 108

Db 107 TAATTTATTAATAATGATATAATAGTTTAAATAATTTAATACTTAATAATTATAATTAA 48

Oy 524 TAAACTGAATCTTCAGCNGAATAACCTTAATATACTTTGTTAG 567

Db 47 AAAGTTATATTAATCTTTATAAATTAAATTATATATATTAG 4

US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
: GENERAL INFORMATION:

APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.

; APPLICANT: Wellem, Thomas E.
 ; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 ; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEIN
 ; NUMBER OF SEQUENCES: 1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

1 CITY: Newport Beach
2 STATE: California
3 COUNTRY: US
4 ZIP: 92660
5
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patentn Release #1.0, Version #1.25
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/487,826B
13 FILING DATE: 10-SEP-1993
14 CLASSIFICATION: 435
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Israelsen, Ned
17 REGISTRATION NUMBER: 29,655
18 REFERENCE/DOCKET NUMBER: NIH121.001CP1
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (619) 235-8550
21 TELEFAX: (619) 235-0176
22 INFORMATION FOR SEQ ID NO: 13:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 19124 base pairs
25 TYPE: nucleic acid
26 STRANDEDNESS: single
27 TOPOLOGY: linear
28 MOLECULE TYPE: cDNA
29 HYPOTHETICAL: NO
30 ANTI-SENSE: NO
31
32 US-08-487-826B-13

```

Query Match	7.1%;	Score 48.6;	DB 2;	Length 19124;
Best Local Similarity	45.7%;	Pred. No. 0.14;		
Matches 200;	Conservative	0;	Mismatches 237;	Indels 1;
				Gaps 1

Db 1877 TTAATCGATGACATAAAATTATATTATATTAGAAATTAATAATTCAAAATTATAAAAAAT 1933

Db 1937 ATGGAATGTTTGTATATTATTITTTTAAAAAT -TTAATTATTATTATTATTATTTA 1995

Db 1996 TTTT TTTT TTTT TTTT TGTG TCTA AATA AAAA AGCA ATATG ATTCAG TAAAAA ATATAT 2055

Qy 315 TGTG TTCA AGCTCG ATCATAT TCTTT TATTG AATTGG GAAGAA AATAC TGTTC TG 374

Db 2056 ATATTTACATTAATGGCAAAATAATTGTTATTATTATTATATGACTATAATAATATTTAG 2115

Oy 375 ATAGCATGAAATGCAAAATTTTATCTCNCATAATTTTAAGAACTATTGAGA 434

DB 2116 ATTAAACATGTAAATTCATTTAACAGAAATAAATAATTATATATATATTAATT 2179

DY 435 AATTGATTAATGCATGA 452

Search completed: November 27, 2003, 12:32:17

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:02:42 ; Search time 4026.04 Seconds
(without alignments)
10100.295 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 994
Sequence: 1 ctcaccagctgtctctccag.....ccacctggggtggtggat 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_ey:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	978	98.4	994	6	ARI46581	ARI46581 Sequence
2	978	98.4	994	6	BD079830	BD079830 Cancer-as
3	765.8	77.0	2572	9	HUMSCP2A	M75883 Human stero
4	646.8	65.1	2661	4	AF051897	AF051897 Oryctolag
5	576.8	58.0	1836	10	RATNSLTP	M58287 Rat non-spe
6	576.8	58.0	1945	10	RNNSLIPTP	X60654 Rat mRNA fo
7	576.6	58.0	1980	10	BC034613	BC034613 Mus muscu
8	576.6	58.0	2626	10	BC018384	BC018384 Mus muscu
9	575.2	57.9	2571	10	RATSCPXA	M57453 Rat sterol
10	575.2	57.9	2599	10	RAT60KDA	M62763 Rat 60 kDa
11	573.4	57.7	2152	6	AX306091	AX306091 Sequence
12	573.4	57.7	2152	10	MUSSCP	M91458 Mus musculu
13	522.2	52.5	1409	10	RATSCP2	M34728 Rat sterol
14	508.2	51.1	1500	9	HUMSCP2A	M75884 Human stero
15	425.6	42.8	2443	5	CHKSCP2A	L09231 Gallus gall
16	280.2	28.2	1359	3	D86473	D86473 Caenorhabdi
17	267.6	26.9	1916	3	AY089539	AY089539 Drosophi
18	266	26.8	1375	3	AY058779	AY058779 Drosophi
19	248.6	25.0	30171	3	CEY57A10C	AL023847 Caenorhab
20	243	24.4	1376	3	DMSCPXRTTH	X95443 D.melanogab
21	243	24.4	1804	3	DMSCPX	X97685 D.melanogab
22	241.4	24.3	87634	2	CEK01F6	Z81559 Caenorhabdi
23	214	21.5	56414	1	MTV002	AL008967 Mycobacte
24	214	21.5	307550	1	BX248343	BX248343 Mycobacte
25	211.4	21.3	167777	3	AC009204	AC009204 Drosophi
26	211.4	21.3	172061	3	AC009203	AC009203 Drosophi
27	211.4	21.3	232737	3	AE003660	AE003660 Drosophi
28	208.2	20.9	85699	2	AC020021	AC020021 Drosophi
29	208.2	20.9	167777	3	AC009204	AC009204 Drosophi
30	208.2	20.9	172061	3	AC009203	AC009203 Drosophi
31	208.2	20.9	232737	3	AE003660	AE003660 Drosophi
32	203	20.4	128469	2	AC020029	AC020029 Drosophi
33	199.6	20.1	14995	1	AE007112	AE007112 Mycobacte
34	196.2	19.7	300425	1	AP005044	AP005044 Streptomy
35	184.4	18.6	103284	8	NCB9B11	BX284747 Neurospor
36	161.6	16.3	785	10	MUSSCPM	M62361 Mus musculu
37	154.4	15.5	201	9	HSSCPXT09	U11306 Human stero
38	154.4	15.5	193774	9	AL445183	AL445183 Human DNA
39	145.2	14.6	198	9	HSSCPXT10	U11307 Human stero
40	132.6	13.3	40281	1	MLCB22	Z98741 Mycobacteri
41	132.6	13.3	348450	1	MLEPRTN4	AL583920 Mycobacte
42	131.2	13.2	2381	1	AF306830	AF306830 Streptomy
43	125.4	12.6	13287	1	AE011116	AE011116 Methanosa
44	119.2	12.0	1725	9	HSSCPXT12	U11309 Human stero
45	118	11.9	248677	2	AL844206	AL844206 Mus muscu

ALIGNMENTS

RESULT 1
ARI46581
LOCUS ARI46581 994 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 21 from patent US 6218521.
ACCESSION ARI46581
VERSION ARI46581.1 GI:15109770
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 994)
AUTHORS Obata,Y.
TITLE Isolated nucleic acid molecules associated with gastric cancer and
methods for diagnosing and treating gastric cancer
JOURNAL Patent: US 6218521-A 21 17-APR-2001;

FEATURES		Location/Qualifiers	
source		1..994	
		/organism="unknown"	
BASE COUNT	313 a	168 c	249 g 248 t 16 others
ORIGIN			
Query Match 98.4%; Score 978; DB 6; Length 994;			
Best Local Similarity 100.0%; Pred. No. 4.2e-234;			
Matches 994; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CTCACCACAGTTGCTCTCAGATGTTGGGTATGCTGGAAGAACATATGAAAAATATG	60
DB	1	CTCACCACAGTTGCTCTCAGATGTTGGGTATGCTGGAAGAACATATGAAAAATATG	60
QY	61	GAACAAAAATTGAACACTTTGCCAAAAATTGGATGAAAAATCATTAACATTCAGTTAATA	120
DB	61	GAACAAAAATTGAACACTTTGCCAAAAATTGGATGAAAAATCATTAACATTCAGTTAATA	120
QY	121	ACCCGTATTCAGTTCACAAGATGAATACAGTTAGATGAAGTGAAGCATCTAAAGAAG	180
DB	121	ACCCGTATTCAGTTCACAAGATGAATACAGTTAGATGAAGTGAAGCATCTAAAGAAG	180
QY	181	TTTTGATTTTTTGACTATCTTACAATGTTGTCCCACTTCAGATGGTGTGCAGACGAA	240
DB	181	TTTTGATTTTTTGACTATCTTACAATGTTGTCCCACTTCAGATGGTGTGCAGACGAA	240
QY	241	TTTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCCTGCAATCCAAAGCTGTGAAATTT	300
DB	241	TTTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCCTGCAATCCAAAGCTGTGAAATTT	300
QY	301	TGGCACAAGAATGATGACTGATTTGCCAAGCTCGTTGAAGAAAAAGCATTATTAAAA	360
DB	301	TGGCACAAGAATGATGACTGATTTGCCAAGCTCGTTGAAGAAAAAGCATTATTAAAA	360
QY	361	TGGTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAAATGCTATGAGAAATCTGGCCTGA	420
DB	361	TGGTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAAATGCTATGAGAAATCTGGCCTGA	420
QY	421	CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCCTTA	480
DB	421	CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCCTTA	480
QY	481	CTTATGAAGCACTGGGACTCTGTCCAGAAGACAAAGTGCAACGCTGCTGTGATAGAGAG	540
DB	481	CTTATGAAGCACTGGGACTCTGTCCAGAAGACAAAGTGCAACGCTGCTGTGATAGAGAG	540
QY	541	ATAATACATATGAGGAAAGTGGGTCAATAATCCTAGTGTGACTGATTTCAAAGGGAC	600
DB	541	ATAATACATATGAGGAAAGTGGGTCAATAATCCTAGTGTGACTGATTTCAAAGGGAC	600
QY	601	ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGGG	660
DB	601	ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGGG	660
QY	661	AAGCCGAAAAAGAGGCAAAGTTCCTGCTGCAAAAGTGCTCTGNCATATTTANGCAT	720
DB	661	AAGCCGAAAAAGAGGCAAAGTTCCTGCTGCAAAAGTGCTCTGNCATATTTANGCAT	720
QY	721	TGAGGAACCTGTGTTGTAACACTCTACAAGATGGGGTTTTCCCGGAAGCCGCAAGTTC	780
DB	721	TGAGGAACCTGTGTTGTAACACTCTACAAGATGGGGTTTTCCCGGAAGCCGCAAGTTC	780
QY	781	TTTATGAACCTCATCAAAATTTGAAGCNGTTCCAACCAAGCTCTGCAAGTNATNGTTAA	840
DB	781	TTTATGAACCTCATCAAAATTTGAAGCNGTTCCAACCAAGCTCTGCAAGTNATNGTTAA	840
QY	841	NGNAAAATCTNGTTTAAAGNGGATTTGAGAAGGAAACNTNAAAGAGGGAANGGGAACA	900
DB	841	NGNAAAATCTNGTTTAAAGNGGATTTGAGAAGGAAACNTNAAAGAGGGAANGGGAACA	900
QY	901	ATTTGTGAAGAAAAATNCGNGGGAATTTTGGCCCTTCAAGGGGAAANAATGCGCCTGG	960
DB	901	ATTTGTGAAGAAAAATNCGNGGGAATTTTGGCCCTTCAAGGGGAAANAATGCGCCTGG	960

QY	961	GGGGTAAAGANGGCCACCTTGGGGTGTGGAT	994	
Db	961	GGGGTAAAGANGGCCACCTTGGGGTGTGGAT	994	
RESULT 2				
BD079830 994 bp DNA linear PAT 27-AUG-2002				
LOCUS BD079830				
DEFINITION Cancer-associated nucleic acids and polypeptides.				
ACCESSION BD079830				
VERSION BD079830.1 GI:22625433				
KEYWORDS JP 2001516009-A/496.				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE 1 (bases 1 to 994)				
AUTHORS Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I., Oghare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.				
TITLE Cancer-associated nucleic acids and polypeptides				
JOURNAL Patent: JP 2001516009-A 496 25-SEP-2001;				
COMMENT LUDWIG INSTITUTE FOR CANCER RESEARCH				
OS Homo sapiens (human)				
PN JP 2001516009-A/496				
PD 25-SEP-2001				
PF 15-JUL-1998 JP 2000503425				
PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR				
10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR				
11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD				
J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG				
CHEN,				
PI IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI				
OZLEM TURECI,				
PI UGUR SAHIN				
PC G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC				
A61P35/00,				
PC C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,				
PC C12N15/00				
CC Cancer-associated nucleic acids and polypeptides. FH Key				
Location/Qualifiers				
FT source 1. .994				
FT /organism='Homo sapiens (human)'				
FEATURES				
source				
1. .994				
/organism="Homo sapiens"				
/mol_type="genomic DNA"				
/db_xref="taxon:9606"				
BASE COUNT	313 a	168 c	249 g	248 t 16 others
ORIGIN				
Query Match 98.4%; Score 978; DB 6; Length 994;				
Best Local Similarity 100.0%; Pred. No. 4.2e-234;				
Matches 994; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	CTCACCAGTTGCTCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG	60	
Db	1	CTCACCAGTTGCTCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG	60	
QY	61	GAACAAAAATTGAACACTTTGCCAAAAATTGGATGGAATAATCATTAACATTCAGTTAATA	120	
Db	61	GAACAAAAATTGAACACTTTGCCAAAAATTGGATGGAATAATCATTAACATTCAGTTAATA	120	
QY	121	ACCCGTATCCAGTTCACAAGATGAATACAGTTAGATGAAGTGAAGCATCTTAAGAAG	180	
Db	121	ACCCGTATCCAGTTCACAAGATGAATACAGTTAGATGAAGTGAAGCATCTTAAGAAG	180	
QY	181	TTTTGATTTTTTGACTATCTTACATGTTGTCCCACTTCAGATGCTGCAGACGAA	240	
Db	181	TTTTGATTTTTTGACTATCTTACATGTTGTCCCACTTCAGATGCTGCAGACGAA	240	
QY	241	TTTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCCTGCAATCCAAAGCTGTGAAATTT	300	

Db 241 TTTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCCCTGCATCCAAAGCTGTGAAATTT 300

Qy 301 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTA 360

Db 301 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTA 360

Qy 361 TGGTTGGCTTTGATATGAGTAAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGA 420

Db 361 TGGTTGGCTTTGATATGAGTAAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGA 420

Qy 421 CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCTTA 480

Db 421 CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCTTA 480

Qy 481 CTTATGAAGCACTGGGACTCTGTCCAGAAGACAGAGTGCAAGCTGGTTGATAGAGAG 540

Db 481 CTTATGAAGCACTGGGACTCTGTCCAGAAGACAGAGTGCAAGCTGGTTGATAGAGAG 540

Qy 541 ATAATACATATGAGAGAAAGTGGTCTATAATCCTAGTGTGACTGATTTCCAAAGGAG 600

Db 541 ATAATACATATGAGAGAAAGTGGTCTATAATCCTAGTGTGACTGATTTCCAAAGGAG 600

Qy 601 ACCCACTAGGCGCTACAGGCTCTGCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGG 660

Db 601 ACCCACTAGGCGCTACAGGCTCTGCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGG 660

Qy 661 AAGCCGGAAGAGAGCAAGTTCCTGCTGCAAGAGTGGCTCTGCNGCATTAATTANGCAT 720

Db 661 AAGCCGGAAGAGAGCAAGTTCCTGCTGCAAGAGTGGCTCTGCNGCATTAATTANGCAT 720

Qy 721 TGGAGGAATCTGTGTTGTAACACTCTACAAGATGGGGTTTCCCGGAAGCCGCACTTCC 780

Db 721 TGGAGGAATCTGTGTTGTAACACTCTACAAGATGGGGTTTCCCGGAAGCCGCACTTCC 780

Qy 781 TTTTAGAACTCATCAAAATTGAAGCCNGTTCCCAACCAAGCTCTGCAAGTATNGGTTAA 840

Db 781 TTTTAGAACTCATCAAAATTGAAGCCNGTTCCCAACCAAGCTCTGCAAGTATNGGTTAA 840

Qy 841 NGNAAAATCTNGTTTAAAGNGAGATTGAGAAGGAAACNTNAAAAGGGAAGGGGAACA 900

Db 841 NGNAAAATCTNGTTTAAAGNGAGATTGAGAAGGAAACNTNAAAAGGGAAGGGGAACA 900

Qy 901 ATTTGTGAAAAGAAAAATNCGNGGGAATTTTGGCCCTTCAAGGGAAAAAATGAGCCCTGG 960

Db 901 ATTTGTGAAAAGAAAAATNCGNGGGAATTTTGGCCCTTCAAGGGAAAAAATGAGCCCTGG 960

Qy 961 GGGGTAAAGANGGCCACCCCTGGGCTGGGGAT 994

Db 961 GGGGTAAAGANGGCCACCCCTGGGCTGGGGAT 994

RESULT 3
HUMSCP2A 2572 bp mRNA linear PRI 06-DEC-1993
LOCUS Human sterol carrier protein X/sterol carrier protein 2 mRNA,
DEFINITION complete cds.
ACCESSION M75883
VERSION M75883.1 GI:432974
KEYWORDS sterol carrier protein-2, sterol carrier protein X.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2572)
AUTHORS He,Z., Yamamoto,R., Furch,E.E., Schantz,L.J., Naylor,S.L.,
TITLE George,H., Billheimer,J.T. and Straus,J.F. III.
CDNAs encoding members of a family of proteins related to human
sterol carrier protein 2 and assignment of the gene to human
chromosome 1 p21---pter
JOURNAL DNA Cell Biol. 10 (8), 559-569 (1991)
MEDLINE 92029618
PUBMED 1718316

REFERENCE 2 (bases 1 to 2572)
AUTHORS Vesa,J., Hellsten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T.,
Mead,S., Cowell,J.K., Straus,J.F.III. and Peltonen,L.
TITLE Assignment of sterol carrier protein X/sterol carrier protein 2 to
1p32 and exclusion as the causative gene for infantile neuronal
ceoid lipofusiosis
JOURNAL Unpublished
COMMENT On Dec 6, 1993 this sequence version replaced gi:410029.
ORIGINAL source text: Human liver cDNA to mRNA.
FEATURES
source
1. .2572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="1p32"
/tissue_type="liver"
/clone_lib="Clontech HL-1001b;HL-1115b"
1. .2572
/gene="SCP-X/SCP-2"
22. .1665
/gene="SCP-X/SCP-2"
/codon_start=1
/product="sterol carrier protein X/sterol carrier
protein-2"
/protein_id="AAA03557.1"
/db_xref="GI:432975"
/translation="MSSSPWBPATLRVFGVGMKTFVKGPAENSRDYPDLAEAGK
KALADAQIPYSAVDOACVGVFGDSTCGORAIYHSLGMTGIPINNNNCATGSTALL
MARQLIQGVAEVCLALGFPEKMSKSLGKISDRITPTDKHVDLLINKYGLSAHPVAP
QMFYAGKEHMEKYGKIEHFAKIGKMHKSNNPYSOFODEYSLDEVMAKVEFDF
LTLQCCPTSDGAAAAILASBAFVQKYGLQSKAVEILAOEMTDLPSFEEKSIIMV
GFDMSKEARKCYEKSGLTPNDIDVIELHDCFSNELLTYBALGLCPREGGATLVDRG
DNTYGGKWVINSGLISKGPLGATGLAOCAELCQLRGEAGKROVPGAKVALQHL
GIGAVVVTLYKMGFPEAASSFRTHQIEAVPTSSASDGFKANLVFKEIEKLEEGEO
FVKKIGIFAFKVDGPGGKEATWVDVKNQGSVLPSNDKADCTITMADSDFLALM
TGKMPQSAFFPGKLUKITGNMGLAMKLQNLQLOPGNAKL"
1932. .1937
/gene="SCP-X/SCP-2"
/evidence=not_experimental
polyA_signal
2213. .2218
/gene="SCP-X/SCP-2"
/evidence=not_experimental
polyA_signal
2258. .2263
/gene="SCP-X/SCP-2"
/evidence=not_experimental
polyA_signal
2341. .2346
/gene="SCP-X/SCP-2"
/evidence=experimental
polyA_signal
2550. .2555
/gene="SCP-X/SCP-2"
/evidence=experimental
polyA_site
2572
/gene="SCP-X/SCP-2"
/evidence=experimental
BASE COUNT 794 a 439 c 570 g 769 t
ORIGIN
Query Match 77.0%; Score 765.8; DB 9; Length 2572;
Best Local Similarity 91.1%; Pred. No. 5.9e-181;
Matches 897; Conservative 0; Mismatches 68; Indels 20; Gaps 8;
Qy 1 CTCACCCAGTTGCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 60
Db 485 CTCACCCAGTTGCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 544
Qy 61 GAACAAAATATGAACACTTGGCAAAAATTTGGATGGAAGAAATCATTAACATTCAGTTAATA 120
Db 545 GAACAAAATATGAACACTTGGCAAAAATTTGGATGGAAGAAATCATTAACATTCAGTTAATA 604
Qy 121 ACCCGTATTCCTCAGTTCAGATGAATACAGTTTAGATGAAGTGTGCGATCTAAAGAG 180
Db 605 ACCCGTATTCCTCAGTTCAGATGAATACAGTTTAGATGAAGTGTGCGATCTAAAGAG 664

QY	181	TTTTGATTTTGTGACTATCTTACAGTGTGTCCCACTTCAGATGCTGTCAGACGAA	240
Db	665	TTTTTGAATTTTGTGACTATCTTACAGTGTGTCCCACTTCAGATGCTGTCAGACGAA	724
QY	241	TTTTGGCCAGTGAAGCATTGTGTACAGATATGGCCTGCAATCCAAAGCTGTGGAATTT	300
Db	725	TTTTGGCCAGTGAAGCATTGTGTACAGATATGGCCTGCAATCCAAAGCTGTGGAATTT	784
QY	301	TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAGAAAAAGCATTATTA	360
Db	785	TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAGAAAAAGCATTATTA	844
QY	361	TGTTGGCTTTGATATGAGTAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA	420
Db	845	TGTTGGCTTTGATATGAGTAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA	904
QY	421	CACCAATGATATTTGACGTAATAGAACTTCAAGATTGCTTTTCTAACCAACGACTCTTA	480
Db	905	CACCAATGATATTTGACGTAATAGAACTTCAAGATTGCTTTTCTAACCAACGACTCTTA	964
QY	481	CTTATGAACACTGGGACTCTGTCCAGAGACAGAGTGCAACGCTGCTGATAGAGAG	540
Db	965	CTTATGAACACTGGGACTCTGTCCAGAGACAGAGTGCAACGCTGCTGATAGAGAG	1024
QY	541	ATAATACATATGAGAGAAAGTGGGTATAATCTAGTGTGACTGATTTCAAAGGAC	600
Db	1025	ATAATACATATGAGAGAAAGTGGGTATAATCTAGTGTGACTGATTTCAAAGGAC	1084
QY	601	ACCCACTAGGCGCTACAGGCTTGTCTAGTGTGACAGAACTCTGTGGCAGCTGAGAGGG	660
Db	1085	ACCCACTAGGCGCTACAGGCTTGTCTAGTGTGACAGAACTCTGTGGCAGCTGAGAGGG	1144
QY	661	AAGCCGAAAAAGGCAAGTTCCTGCTGCAAAAGGTGCTCTGCGCATTAATTTANGCAT	720
Db	1145	AAGCCGAAAA--GAGCAAGTTCCTGCTGCAAAAGGTGCTCTGCGCATTAATTTANGCAT	1202
QY	721	TGAGGAACTGTGTTGTAACACTCTCAAGATGGGTTTCCCGAAGCCGCACTTCC	780
Db	1203	TGAGGAACTGTGTTGTAACACTCTCAAGATGGG--TTTCCGGAAGCCGCACTT--C	1259
QY	781	TTTTGAACATCATCAAAATTTGAAGCNGTTCACCAAGCTCTGCAAGTNATNGTTTAA	840
Db	1260	TTTTGAACATCATCAAAATTTGAAGCNGTTCACCAAGCTCTGCAAGTNATNGTTTAA	1316
QY	841	NGNAAATCTNGTTTTAAAGNGATTTGAGAGGAAACNTNAAAGGGAANGGGAACA	900
Db	1317	GGCAAAATCTGTTTAA--GGAGATTGAGAAAGAACTTGA---AGAGGAAGGGAACA	1370
QY	901	ATTGTGAAAGAAAAATNCGGNGGAATTTTCCCTTCAAGGGGAAANAATGGCCCTGG	960
Db	1371	GTTTGTGAAGAAATCGGTGT-----ATTTTGCCCTTCAAGGTGAAG-ATGGCCCTGG	1424
QY	961	GGGTAAGAGANGGCCACCTGGGG	985
Db	1425	GGGTAAGAGANGGCCACCTGGGTGTG	1449
RESULT 4			
LOCUS	AF051897	2661 bp	mRNA linear MAM 30-APR-1998
DEFINITION	Oryctolagus cuniculus sterol carrier protein X (SCP2) mRNA, complete cds.		
ACCESSION	AF051897		
VERSION	AF051897.1	GI:3093801	
KEYWORDS	Oryctolagus cuniculus (rabbit)		
SOURCE	Oryctolagus cuniculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
REFERENCE	1 (bases 1 to 2661)		
AUTHORS	Weber, F.E., Werder, M., Boffelli, D. and Hauser, H.		
TITLE	In pre-sterol carrier protein 2 (SCP2) in solution the leader peptide 1-20 is flexibly disordered and the residues 21-143 adopt		

JOURNAL	the same globular fold as in mature SCP2		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 2661)		
TITLE	Weber, F.E., Werder, M., Boffelli, D. and Hauser, H.		
JOURNAL	Direct Submission		
	Submitted (02-MAR-1998) Kieferchirurgie, Universitaetspital, Raemistr. 100, Zurich 8091, Switzerland		
FEATURES	location/Qualifiers		
source	1..2661		
	/organism="Oryctolagus cuniculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9986"		
gene	/cell_type="enterocytes"		
	1..2661		
	/gene="SCP2"		
CDS	53..1696		
	/gene="SCP2"		
	/note="peroxisomal 3-oxoacyl coenzyme A thiolase"		
	/codon_start=1		
	/product="sterol carrier protein X"		
	/protein_id="AAC15422.1"		
	/db_xref="GI:3093802"		
	/translation="MSSSARKLAPLRVAVGVGMTKFKPGTEDARDYPDMAKEAGQ KALADAOIPYSAVEQACIGYVGDSTGORA VYHSLGTGIPINVNNCTGSTALF MGRQLIQGMACEVLALGFEMERGLGAKFPDRTPMDKHLDVLINXGLSAHPVAP OMFGSAGKEHMEKYGTKIEHFAKIGKXNHSHSVNPNPYSOFQKEYSLDEWMSRPIDF LTVLQCCPTSDGAAAAIILASEEFVKRYGLOSKAVEILAOEMVTFPSSFEKSI IKMV GFDMSKEARRCYEKSGLRPSDIDVIELHDCFSANELTYEALGLCEPGKGLVDRG DNTYGGKWINPSGGLISKHPLGATGLAQCAELCWLQRGBAGKQYPGAQVALQHLN GLGGAAVTLVKMGFPPEAASSFRTHQIEAAPSSAGDFKANLVFKEIEKLEEGEQ FVKIGGI FAFKVKDGPGGKEATWVDVKNKGSVLPNSDKADCTITIASDILLALM TGKMPQSAFPGKLIKITGNMGLAMKLNQLQPGKAKL"		
BASE COUNT	806 a 490 c 623 g 742 t		
ORIGIN			
Query Match	65.1%; Score 646.8; DB 4; Length 2661;		
Best Local Similarity	88.0%; Pred. No. 3.8e-151;		
Matches 757; Conservative	0; Mismatches 95; Indels 8; Gaps 5;		
QY	1	CTCACCAGTGTCTCCTCAGATGTTGGGTATGCTGGAAGAACAATATGAAAAATATG	60
Db	516	CCCACCAGTGTCTCCTCAGATGTTGGGTCTGCGAAAAAGACACATGAAAAATATG	575
QY	61	GAACAAAAATTTGAACACTTTGCAAAAAATTTGATGGAAGAAATCATAACATTCACTTAATA	120
Db	576	GAACGAAAAATTTGAACACTTTGCAAAAAATTTGGTGGAAAAACCATTAAGCATTCAGTTAATA	635
QY	121	ACCCGATATCCCAAGTTCAGAGATGAATACAGTTTGAAGAGTGATGGCATCTTAAAGAAG	180
Db	636	ACCCGATATCCCAAGTTCAGAAAGAAATACAGTTTGAAGAGTGATGGCATCTTAAAGCA	695
QY	181	TTTTGATTTTGTGACTATCTTACAAATGTTGCCACTTCAGATGCTGTCAGACGCAA	240
Db	696	TTTTGATTTTGTGACTGCTTACAGTCTGTCCCACTTCAGATGCTGTCAGACGCAA	755
QY	241	TTTGGCCAGTGAAGCATTTGTACAGAGATATGGCCTGCAATCCAAAGCTGTGAAAAATT	300
Db	756	TTTTGGCCAGTGAAGATTTGTGAAGAGATATGGCTTGCAAGTCCAAAGCTGTGAAAAATT	815
QY	301	TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAGAAAAAAGCATTATTA	360
Db	816	TGGCACAAGAGATGTTAAGTATTTCCAAAGCTCATTGGAAGAAAAAGCATTATTA	875
QY	361	TGTTGGCTTTGATATGAGTAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA	420
Db	876	TGTTGGCTTTGATATGAGTAAGAGCTGCCAGAAAGTCTATGAGAAATCTGGCTGA	935
QY	421	CACCAATGATATTTGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAACTCCTTA	480
Db	936	GACCAAGTATATTTGATGTAATAGAACTTCATGATTTTCTGCCAACGAACTCCTCA	995
QY	481	CTTATGAAGCACTGGGACTCTGTCCAGAGGACAAAGGTGCAACGCTGTTGATAGAGAG	540

Db 996 CTTATGAACTAGAGACTCTGTCCGGAAGCAAAAGGTGAGCACTGGTTGACAGAGAG 1055

Qy 541 ATAAATACATATGAGAAAGTGGGTATATAATCCTAGTGTGAGCTGATTTCAAAGGAC 600

Db 1056 ATAAATACTTACGAGGAAAGTGGGTATATAACCTAGTGTGAGATTGATTTCAAAGGAC 1115

Qy 601 ACCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAGAACTCTGTGGCAGCTGAGAGGG 660

Db 1116 ACCCCCTGGAGCTACAGGTCTGTCTCAATGTGCCGAAGCTGTGGCAGCTGAGAGGG 1175

Qy 661 AAGCCGGAAGAGAGCAAGTCTGTGTGCAAGAGTGGCTGTGCGCATTAATTANGCAT 720

Db 1176 AAGCCGGAAG--GAGGCAAGTCTGTGTGCAAGAGTGGCTGTGCGCATTAAGTGGCCT 1233

Qy 721 TGGAGGAAGCTGTGTGTGAACCTCTCAAGATGGGGTTTCCCGGAAGCCGCCAGTTCC 780

Db 1234 CGAGGAGCTGCGGTGTACACTCTCAAGATGGG--TTTCTGTAGGCTGCCAGTT-C 1290

Qy 781 TTTAGAACTCATCAAAATTTGAAGCCNGTTCACCAAGCTCTGCAAGTNATNGTTAA 840

Db 1291 TTTAGAAACCATCAAA--TTGAGGCTGTCCAAAC-AGCTCTGAGGTGATGATTTAA 1347

Qy 841 NGNAAATCTNGTTTAAAG 860

Db 1348 GGCATAATCTGTCTTTAAGG 1367

RESULT 5

RATNSLTP

LOCUS RATNSLTP 1836 bp mRNA linear ROD 27-APR-1993

DEFINITION Rat non-specific lipid transfer protein (nsl-TP) mRNA, 3' end.

ACCESSION M58287

VERSION M58287.1 GI:205769

KEYWORDS non-specific lipid transfer protein.

SOURCE Rattus sp.

ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1836)
Ossendrop, B.C., van Heusden, G.P. and Wirtz, K.W.
The amino acid sequence of rat liver non-specific lipid transfer protein (sterol carrier protein 2) is present in a high molecular weight protein: evidence from cDNA analysis
Biochem. Biophys. Res. Commun. 168 (2), 631-636 (1990)

JOURNAL MEDLINE 90241231

PUBMED 2334427

COMMENT Original source text: Rat liver, cDNA to mRNA.

FEATURES

Source

1..1836
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
/class_type="liver"
1..1836
/gene="nsl-TP"
/gene="nsl-TP"
/codon_start=1
/product="non-specific lipid transfer protein"
/protein_id="AAA1726.1"
/db_xref="GI:205770"

gene

CDS

BASE COUNT 500 a 392 c 503 g 441 t

Qy 541 ATAAATACATATGAGAAAGTGGGTATATAATCCTAGTGTGAGCTGATTTCAAAGGAC 600

Db 1056 ATAAATACTTACGAGGAAAGTGGGTATATAACCTAGTGTGAGATTGATTTCAAAGGAC 1115

Qy 601 ACCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAGAACTCTGTGGCAGCTGAGAGGG 660

Db 1116 ACCCCCTGGAGCTACAGGTCTGTCTCAATGTGCCGAAGCTGTGGCAGCTGAGAGGG 1175

Qy 661 AAGCCGGAAGAGAGCAAGTCTGTGTGCAAGAGTGGCTGTGCGCATTAATTANGCAT 720

Db 1176 AAGCCGGAAG--GAGGCAAGTCTGTGTGCAAGAGTGGCTGTGCGCATTAAGTGGCCT 1233

Qy 721 TGGAGGAAGCTGTGTGTGAACCTCTCAAGATGGGGTTTCCCGGAAGCCGCCAGTTCC 780

Db 1234 CGAGGAGCTGCGGTGTACACTCTCAAGATGGG--TTTCTGTAGGCTGCCAGTT-C 1290

Qy 781 TTTAGAACTCATCAAAATTTGAAGCCNGTTCACCAAGCTCTGCAAGTNATNGTTAA 840

Db 1291 TTTAGAAACCATCAAA--TTGAGGCTGTCCAAAC-AGCTCTGAGGTGATGATTTAA 1347

Qy 841 NGNAAATCTNGTTTAAAG 860

Db 1348 GGCATAATCTGTCTTTAAGG 1367

Qy 541 ATAAATACATATGAGAAAGTGGGTATATAATCCTAGTGTGAGCTGATTTCAAAGGAC 600

Db 1056 ATAAATACTTACGAGGAAAGTGGGTATATAACCTAGTGTGAGATTGATTTCAAAGGAC 1115

Qy 601 ACCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAGAACTCTGTGGCAGCTGAGAGGG 660

Db 1116 ACCCCCTGGAGCTACAGGTCTGTCTCAATGTGCCGAAGCTGTGGCAGCTGAGAGGG 1175

Qy 661 AAGCCGGAAGAGAGCAAGTCTGTGTGCAAGAGTGGCTGTGCGCATTAATTANGCAT 720

Db 1176 AAGCCGGAAG--GAGGCAAGTCTGTGTGCAAGAGTGGCTGTGCGCATTAAGTGGCCT 1233

Qy 721 TGGAGGAAGCTGTGTGTGAACCTCTCAAGATGGGGTTTCCCGGAAGCCGCCAGTTCC 780

Db 1234 CGAGGAGCTGCGGTGTACACTCTCAAGATGGG--TTTCTGTAGGCTGCCAGTT-C 1290

Qy 781 TTTAGAACTCATCAAAATTTGAAGCCNGTTCACCAAGCTCTGCAAGTNATNGTTAA 840

Db 1291 TTTAGAAACCATCAAA--TTGAGGCTGTCCAAAC-AGCTCTGAGGTGATGATTTAA 1347

Qy 841 NGNAAATCTNGTTTAAAG 860

Db 1348 GGCATAATCTGTCTTTAAGG 1367

ORIGIN

Query Match 58.0%; Score 576.8; DB 10; Length 1836;
Best Local Similarity 78.7%; Pred. No. 1.3e-133;
Matches 759; Conservative 0; Mismatches 187; Indels 19; Gaps 6;

Qy 6 CCACTGTCTCCTCAGATGTTGGGTATGTCTGGAAGAAACATATGAAAAATATGAAACA 65

Db 373 CCATTGTCTCCTCAGCTGTTGGGAGCGCTGGGAAGACATATGAAACATATGAAACA 432

Qy 66 AAAATTGAAACACTTTGCAAAAATTGGATGGAAAAATCATTAACATTCAAGTTAACCCG 125

Db 433 AAGGTTGAACACTTTGCAAAAATTGGATGAAAAATCATTAACACTCAGTTAATACCCG 492

Qy 126 TATTCAGGTTCCAGATGAATACAGTTAGATGAAGTATGATGATGATGATGATGATG 185

Db 493 TATTCAGGTTCCAGATGAATACAGTTAGATGAATGAATGAATGAATGAATGAATG 552

Qy 186 GATTTTGTGACTATCTTCAATGTTGTGCCACTTCAAGTGTGCTGACAGCAATTTTG 245

Db 553 GATTTTGTGACTGCTTCAATGCTGTGCCACTCAAGTGTGCTGACAGCAATTTTG 612

Qy 246 GCCAGTGAAGCATTTGTACAGATATGCGCTGCAATCCAAAGCTGTGAAATTTGGCA 305

Db 613 TCTAGTGAAGAGTTGTGTGAGAGATGCGCTGCAATCCAAAGCTGTGAAATTTGGCA 672

Qy 306 CAAGAAATGATGATGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAAGTGTG 365

Db 673 CAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 732

Qy 366 GCGTTGATATGATGAAGAAAGCTGCAAGAAATGCTATGAGAAATCTGGCTGACACA 425

Db 733 GCGTATGATATGATGAAGAAAGCTGCAAGAAATGCTATGAGAAATCTGGCTGACACA 792

Qy 426 AATGATATGATGATGAAGAAAGCTTCAAGATTTGCTTTTCAACCAAGCACTCTTACTTAT 485

Db 793 AGTATGATGATGATGAAGAAAGCTTCAAGATTTGCTTTTCAACCAAGCACTCTTACTTAT 852

Qy 486 GAAGCACTGGAGCTGTGTCAGAGAGCAAGAGTGCAACGCTGTTGATAGAGAGATTAAT 545

Db 853 GAAGCACTGGAGCTGTGTCAGAGAGCAAGAGTGCAACGCTGTTGATAGAGAGATTAAT 912

Qy 546 ACATATGAGAGAAAGTGGGTATTAATCTAGTGTGACTGATTTCAAAGGACACCCA 605

Db 913 ACTTACGAGAGAAAGTGGGTATTAATCTAGTGTGACTGATTTCAAAGGACACCCA 972

Qy 606 CTAGCGCTACAGGCTCTGCTCAGTGTGCAAGCACTGTGCGCAGCTGAGAGGGGAAGCC 665

Db 973 CTGCGTGCACAGGCTGTGCTCAGTGTGCGCAGGAGCTGTGCGCAGCTGAGAGGGGAAGCC 1032

Qy 666 GGAAGAGAGCAAGTCTCTGCTGCAAGAGTGGCTGTGCGCATTAATTANGCATTTGAG 725

Db 1033 GGAAGAGAGCAAGTCTCTGCTGCAAGAGTGGCTGTGCGCATTAATTANGCATTTGAG 1090

Qy 726 GAAGTGTGTTGTAACACTCTACAAGATGGGTTTCCCGGAAGCCGCAAGTCTCTTTTA 785

Db 1091 GAGCTGCTGTTGTACACCTCTACAAGATGGG--TTTCCCGGAAGCTGCAAGCTCC--TTCA 1147

Qy 786 GAAGTGTGTTGTAAGAGCCNGTTCACCAAGCTCTGCAAGTATNGTTTAANGNAA 845

Db 1148 GAAGTGTGTTGTAAGAGCCNGTTCACCAAGCTCTGCAAGTATNGTTTAANGNAA 1204

Qy 846 AATCTNGTTTAAAGNGATTTGAAGAGCAACNTNAAAGAGGAAGGGGAACAATTG 905

Db 1205 AATCTNGTTTAAAGNGATTTGAAGAGCAACNTNAAAGAGGAAGGGGAACAATTG 1258

Qy 906 TGAAGAGAAATNCGNGGAATTTTGGCCCTTCAAGGGGAANAATGGCCCTGGGGGT 965

Db 1259 TGAAGAGAAATNCGNGGAATTTTGGCCCTTCAAGGGGAANAATGGCCCTGGGGGT 1313

Qy 966 AAAAG 970

Db 1314 AGAAG 1318

RESULT 6
RNNSLIPTR LOCUS 1945 bp mRNA linear ROD 17-FEB-1997
DEFINITION Rat mRNA for 58 kDa protein containing the non-specific lipid transfer protein (sterol carrier protein 2).
X60654
ACCESSION X60654.1 GI:56771
KEYWORDS non-specific lipid transfer protein; sterol carrier protein-2.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1
AUTHORS Billheimer,J.T., Strehl,L.L., Davis,G.L., Strauss,J.F. III and Davis,L.G.
TITLE Characterization of a cDNA encoding rat sterol carrier protein2
JOURNAL DNA Cell Biol. 9 (3), 159-165 (1990)
MEDLINE 90253610
PUBMED 2340090
REFERENCE 2
AUTHORS Seedorf,U. and Assmann,G.
TITLE Cloning, expression, and nucleotide sequence of rat liver sterol carrier protein 2 cDNAs
JOURNAL J. Biol. Chem. 266 (1), 630-636 (1991)
MEDLINE 91093192
PUBMED 1985920
REFERENCE 3
AUTHORS Ossendorp,B.C., Van Heusden,G.P., De Beer,A.L., Bos,K., Schouten,G.L. and Wirtz,K.W.
TITLE Identification of the cDNA clone which encodes the 58-kDa protein containing the amino acid sequence of rat liver non-specific lipid-transfer protein (sterol-carrier protein 2). Homology with rat peroxisomal and mitochondrial 3-oxoacyl-CoA thiolases
JOURNAL Eur. J. Biochem. 201 (1), 233-239 (1991)
MEDLINE 92007881
PUBMED 1915369
REFERENCE 4 (bases 1 to 1945)
AUTHORS Ossendorp,B.C.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1991) B.C. Ossendorp, Centre for Biomembranes and lipid, Enzymology, State, University of Utrecht, Post Box 80054, 3508 TB Utrecht, THE NETHERLANDS
COMMENT See also M58287.
FEATURES
source location/Qualifiers
1. 1945 /organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="sprague-dawley"
/db_xref="taxon:10116"
/sex="Female"
/tissue_type="liver"
/clone_lib="lambda gt11"
/dev_stage="adult"
1. 1930 /evidence=experimental
1. 1642
/note="open reading frame for 58 kDa protein containing the non-specific lipid transfer protein"
/codon_start=2
/protein_id="CAA43061.1"
/db_xref="GI:56773"
/db_xref="SWISS-PROT:P11915"
/translation="PSVALNSPLRVFVVGVMTRKMPGGENSRDYPDLAKEAGQKALADROIPIYSAVEQACGYVYGESTCGORAIYHSLGLTGIPIINNNNSTGSTALFM AQLVQGLLANCVLALGFEKMEKESLGTKYSDRNPLEKHIDVLINKYMSACFPAPQ LFGSAGKEHMETYGTKEVHPAKIKMNHKHSVNNPYSOFODEYSLDEIMKSRVFDL TVLQCCPTSDGAAAIYSSEEFVQKHGLQSKAVEIYAQEMVTDMPTFEEKSVIKMVG YDMSKEARKCYEKSGLGRSDVDVIELHDCFSTNELITYEALGLCPGGGALVDRGD NTYGGKWINPSGGLISKHPDLGATGLAQCAELCWQDRGEAGKRQVPGAKVALQHNLG LGGAAVVTLYRMGFPEAASSFRTHQISAAPTSAGDGFKANLIFKEIEKLEEGEEF

	conflict	VKKIGIFAFKVKDGPGEKEATWVVDVKNKGSVLPDSDKKADCTITMADSDLLAMT GKMPQSAFFQGLKLIAGNMGLAMKLQSLQLPDKAKL"
	33. .34	/citation=[2]
	/replace="cg"	
	792	/citation=[2]
	/replace="c"	
	804. .809	/citation=[1]
	/replace="aaaaa"	
	1211. .1642	/product="non-specific lipid transfer protein"
	/protein_id="CAA43060.1"	
	/db_xref="GI:56772"	
	/db_xref="SWISS-PROT:P11915"	
	/translation="MGFPEAASSFRTHQISAAPTSAGDGFKANLIFKEIEKLEEG EEFVKKIGIFAFKVKDGPGEKEATWVVDVKNKGSVLPDSDKKADCTITMADSDLLA LMTGKMPQSAFFQGLKLIAGNMGLAMKLQSLQLPDKAKL"	
	1271. .1639	
	/product="non-specific lipid transfer protein"	
	1211. .1270	
	/note="non-specific lipid transfer protein presequence"	
	BASE COUNT 531 a 415 c 535 g 464 t	
	ORIGIN	
	Query Match 58.0%; Score 576.8; DB 10; Length 1945;	
	Best Local Similarity 78.7%; Pred. No. 1.3e-133;	
	Matches 759; Conservative 0; Mismatches 187; Indels 19; Gaps	
QY	6 CCAATTGCTCTCAGATGTTGGGTATGCTGGAAGAACAATATGGAACA 65	
DB	467 CCATTGCTCTCAGCTGTTGGAGCGCTGGGAAGACATATGGAACATATGGAACA 526	
QY	66 AAAATTGAACACTTTGCAAAAATTGGATGGA AAAAATCATAAACATTCAATTACCCG 125	
DB	527 AAGTTGAACACTTTGCAAAAATTGGATGGA AAAAATCATAAACACTCAGTTAATACCCG 586	
QY	126 TATTCACAGTTCCAGATGAATACAGTTAGATGAAGTGCATCTAAAGAAGTTT 185	
DB	587 TATTCACAGTTCCAGATGAATACAGTTAGATGAATGAATCAAGGCCAGTTTTC 646	
QY	186 GATTTTGTGACTATCTTACAATGTTGTCCACCTTCAGATGCTGCAGACGAATTTG 245	
DB	647 GATTTTGTGACTGCTTACAATGCTGTCCACCTCAGATGCTGCCGACGACGAATTTG 706	
QY	246 GCCAGTGAAGCATTTGTACAGAAGTATGCGCTGCATCCAAAGCTGTGAAATTTGGCA 305	
DB	707 TCTAGTGAAGAGTTGTGACAGAAGCATGCGCTGCAGTCCAAAGCTGTGAAATTTGGCA 766	
QY	306 CAAGAATGATGACTGATTTGCCAAGCTCGTTTGAGAAAAAAGCATTAATTAATGGTT 365	
DB	767 CAGGAGATGCTGACTGACATGCCAGTACATTTGAAGAAAAAGTTATTAAATGGTT 826	
QY	366 GCGTTGATATGATTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGACACCA 425	
DB	827 GCGTTGATATGATTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGACCC 886	
QY	426 AATGATATTGACGTAATGAACCTTCAGATGCTTTTCTACCAAGCAACTCCTTACTTAT 485	
DB	887 AGTGATGTCGACGTGATGAGCTTCACGATGCTTCTTACCAATGAACCTCCTGACTTAT 946	
QY	486 GAAGCACTGGACTCTGTCCAGAAGACAAAGTGCACGCTGTTGATAGAGAGATAAT 545	
DB	947 GAAGCACTGGGCTCTGTCCAGAAGACAAAGTGCACGCTGTTGATAGAGAGACAAC 1006	
QY	546 ACATATGAGGAAAGTGGTCAATAATCCTAGTGGTGAAGCTGATTTCAAAGGACACCCA 605	
DB	1007 ACTTACGAGGAAAGTGGTCAATAACCTAGTGAAGGCTCATCTCCAAGGACACCCA 1066	
QY	606 CTAGCGCTACAGGTTCTTGTCTCAGTGTGCAAGAACTCTGCTGCAGAGGGGAAAGCC 665	

Db 1067 CTGGGTGCACAGGTCCTGGCTCAGTGC GCGAGCTCTGCTGGCAGCTGAGAGGGAAGCC 1126

QY 666 GGAAGAGAGGCAAGTTCCTGGTGC AAGGTGCTCTGCNGCATATTANGCATTTGAG 725

Db 1127 GGAAG--GAGGAGGTTCTGGGCA AAGGTGCTCTGCAGACAATTAGGCTTGAG 1184

QY 726 GAACTGTGCTGTACACTCTACAGATG GGGTTTCCCGGAAGCCCGCAGTTCCTTTA 785

Db 1185 GAGCTGCTGTGTCACCCTCTACAGAAT GGG--TTTCCCGAAGCTGCCAGCTCC-TTCA 1241

QY 786 GAACTCATCAAAATTGAAGCCNGTTC CCAACCAAGCTCTGCAAGTNAATNGTTAANGNA 845

Db 1242 GAACGCACACAGATTCA--GCTGCTCC CACCAGCTCTGCAGGGGATGATTCAAGGCA 1298

QY 846 AATCTNGTTTAAAGNGGATGAGAAG AAGAAACNTNAAAGGGAANGGGAACAATTG 905

Db 1299 ATCTCATTTTAAAGAA-----ATC GAGAAGAAGCTTGAAGAGGAAGGGAAGATTG 1352

QY 906 TGAAGAAGAAATNCGNGGAAATTTT GCGCCTTCAAGGGGAANAATGCCCTGGGGGT 965

Db 1353 TGAAGAAATCGGTGC-----ATT TTTGCTTCAAGTGAAGATGGCCCCGGGGCAA 1407

QY 966 AAAAG 970

Db 1408 AGAAG 1412

RESULT 7

BC034613 1980 bp mRNA linear ROD 07-AUG-2002

LOCUS BC034613

DEFINITION Mus musculus, steroid carrier protein 2, liver, clone MGC:29961

IMAGE:5123611, mRNA, complete cds.

ACCESSION BC034613

VERSION BC034613.1 GI:21961462

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1980)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyik, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAK Plate: 42 Row: n Column: 7

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

Location/Qualifiers

1..1980

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/map="FVB/N"

/clone="MGC:29961 IMAGE:5123611"

/cissue_type="Liver, normal. 5 month old male mouse."

/clone_1b="NCI CGAP_L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

22..1665

/codon_start=1

/product="steroid carrier protein 2, liver"

/protein_id="AAH34613.1"

/db_xref="GI:21961463"

/db_xref="LocusID:20280"

/translation="MPSVALKSPRLRRVVGWMTKFMKPGGENSRDYPDMAKEAGO KALEDAQIPYSAVEQACVGYVYGDSTSGRAIYHSLGTGPIINVNNCSTGSTALF MAHQIQQGLANCVLALGFEKMERGSIKTRSDRTPTDKHIEVLIDKYGLSAHPITP QMFYAGKEHMEKYCTKVEHFAKIGKXNHGHSVNTYSOFODEYSLSEVMKSPVDF LTILOCCPTSDGAAAAILSSSEFVQYQGLOSKAVEIVAQEMTDLPSFEEKSIKVV GYDMSKEAARCYEKSGLTPTNDVDVIELHDFSVNELITYEALGLCPGOGGTLVDRG DNTYGGKWIINPSGGLISKGPLGATGLAOCAELCWQLRGEAGKROVPKAVLQHNL GLGAVVVTLYRMGFPEAASSFRTHQVSAAPTSAGDGKASLVFKEIEKLEBEGEO FVKKIGGIFAFKVKYKDPGGEKATWVDVKNKGSVLPNSDKADCTITMADSDLLALM TGKMPQSAFPQGLKLIAGNMGLAMKLNQLQPGAKL"

BASE COUNT 556 a 417 c 537 g 470 t

ORIGIN

Query Match 58.0%; Score 576.6; DB 10; Length 1980;

Best Local Similarity 79.1%; Pred. No. 1.4e-133;

Matches 767; Conservative 0; Mismatches 184; Indels 19; Gaps 7;

QY 1 CTCACCCAGTTGCTCTCAGATGTTGGGTATGCTGGAAGAAAGACATATGAAAAATATG 60

Db 485 CACATCCGATTACTCTCAGATGTTGGGTATGCTGGGAAGAAACATATGAAAAATATG 544

QY 61 GAACAAAATTGAACACTTTCAAAAATTGATGAGAAATCATAAACATTCACTTAATA 120

Db 545 GAACAAAAGTTGAACACTTTCAAAAATTGATGAGAAATCATAAACACTCACTTAATA 604

QY 121 ACCCGTATTCACAGTTCAGATGATATACAGTTTAGATGAAGTGATGGCATCTAAGAG 180

Db 605 ACACGTATTCACAGTTCAGATGATATACAGTTTAGATGAAGTGATGGCATCTAAGAG 664

QY 181 TTTTGTATTTTGTACTATCTTACAATGTTGTCCCACTTCAGATGCTGTCAGACAGCA 240

Db 665 TTTTGTATTTTGTACTATCTTACAATGTTGTCCCACTTCAGATGCTGTCAGACAGCA 724

QY 241 TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCTGCAATCCAAAGCTGTGAAATTT 300

Db 725 TTCTGTCAGCAGAGATTGTGTGACAGCAGTACGGCTGCAAGTCCAAAGCGGTGAGATTG 784

QY 301 TGGCACAAGAAATGATGATGATTTGGCCAAGCTCGTTTGAAGAAAAAGCATTTATTA 360

Db 785 TGGCCACAGAGATGATGATGATTTGGCCAAGCTCGTTTGAAGAAAAAGCATTTATTA 844

QY 361 TGGTTGGCTTTGATATAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCGTGA 420

Db 845 TGGTTGGCTATGATATAGTAAGAAGCTGCCAGAGATGCTATGAGAAATCTGGCGTGA 904

QY 421 CACCAATGATATTGACGTAATAGAACTTCACGATTCCTTTTCTACCAACGAATCCTTA 480

Db 905 CACCCACGATGTCGACGTGATAGAGCTTCACGATTCCTTTCTGTCATGAACTCATCA 964

QY 481 CTTATGAAGCACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 540

Db 965 CTTACGAAGCACTGGGCTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 1024

QY 541 ATAATACATATGAGAGAAAGTGGTCATTAATCCTAGTGGTGAAGTGAATTTCAAGGAG 600

Db 1025 ACAACACTTACGAGAGAAAGTGGTCATCAACCTTAGTGAAGGCTCATCTCAAGGAG 1084

QY 601 ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGAGAACTCTGCTGGCAGCTGAGAGGG 660

Db 1085 ACCCACTGGGCGCAACAGGCTGTGCTCAGTGC GCGAGCTCTGCTGGCAGCTGAGAGGG 1144

QY 661 AAGCCGGAAGAGCAAACTTCTCTGCTGCAAGGTGCTCTGCNGCATATTANGCAT 720

Db 1145 AAGCCGGAAA--GAGGCAGGTTCCCGGGCAAAAGTGCGCTGTGACGACAATCTAGGCCT 1202
QY 721 TGGAGAACTGTGGTGTGTAACACTCTACAGATGGGGTTTCCCGAAGCCGCCAGTTCC 780
Db 1203 CGGAGAGCTGTGGTGTGTGTACCCCTCTACAGATGGG--TTTTCGGAAGCTGCCAGCTCC 1260
QY 781 TTTTGAAGCTCATCAAAATTTGAAGCCNGTTCCAACCAAGCTCTGCAGTNGTTTAA 840
Db 1261 TTACAGGA--GCACCAAGTTTCAAGCTGCTCCCA--CCAGCTCTGCAGGGAGTGAATCAA 1316
QY 841 NGNAAATCTNGTTTAAAGNGAATTGAGAAGGAACNTNAAAGGAANGGGAACA 900
Db 1317 GGCAAACTCGTCTTTAA--GGAGATTGAGAAGAAGCTTGA---AGAGGAAGGGAACA 1370
QY 901 ATTGTGAAAGAAAATNCGGNGGAATTTTGGCCCTTCAAGGGGAANAATGGCCCTGG 960
Db 1371 GTTCGTGAAGAAAATCGGTGC-----ATTTTGCCTTCAAGTGAAGAAGATGGCCCTGGA 1425
QY 961 GGGGTAAAAG 970
Db 1426 GGCAAGGAAG 1435

RESULT 8
BC018384
LOCUS BC018384 2626 bp mRNA linear ROD 20-SEP-2002
DEFINITION Mus musculus, sterol carrier protein 2, liver, clone MGC:11534
IMAGE:3710340, mRNA, complete cds.
ACCESSION BC018384
VERSION BC018384.1 GI:17390902
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 17 Row: n Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
Source
1.2626
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11534 IMAGE:3710340"
/tissue_type="Mammary tumor, C3(1)-Tag model, Infiltrating ductal Carcinoma, 5 month old virgin mouse."

/clone_1ib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
56.1699
/codon_start=1
/product="sterol carrier protein 2, liver"
/protein_id="AAH18384.1"
/db_xref="GI:17390903"
/db_xref="LocustID:20280"
/translation="MPSVALKSPRLRRVFGVGMTKFMKPGGENSRDYPDMAKEAGQ
KALEDAQIPYSAVEQACVGYVYDSTSGQRAIYHSLGTGIPINNNNGSTALF
MAHQLIQGLANCVLALGFEKMERGSGITKFSDRTPPTDKHIEVLIDKYGSAHPITP
QMFYAGKEHMERKGTKEVHFAKIGWKNHKSNNNTYSQFODEYSLSEVMSKPVDF
LTILOCCPTSDGAAAILSSSEFVQYGLQSKAVEIVAQEMMTDLPSTFEKSIIVK
GYDMSKEAARCYEKSGLTPNDVDVIELHDFSVNELITYEALGLCPGGGGLVDRG
DNITYGKWINPESGLISKHPLGATGTLAQCAELCWQLRGEAGKROVPKVALQHL
GLGAVVVTLYRMGFPEAASSFRTHOVSAPTSADGFKANLVFKEIEKLEEEGEO
FVKIGGIFAFKYVDGPFGEKEATWVDVKNKGSVLPNSDKKADCTITWADSDLLALM
TGKMPQSAFFQGLKLIAGNMLAMKIQNLQIPGAKL"

BASE COUNT 765 a 521 c 642 g 698 t
ORIGIN
Query Match 58.0%; Score 576.6; DB 10; Length 2626;
Best Local Similarity 79.1%; Pred. No. 1.4e-133;
Matches 767; Conservative 0; Mismatches 184; Indels 19; Gaps 7;

QY 1 CTCACCCAGTGTGCTCCTCAGATGTTGGGTATGCTGGAAGAACAATATG 60
Db 519 CACATCCGATTACTCCTCAGATGTTGGGTATGCTGGAGAAACATATG 578
QY 61 GAACAAAATTGAACACCTTGCAAAAATTGGATGAAAATCATAAACATTCAGTTAATA 120
Db 579 GAACAAAAGTTGAACACCTTGCAAAAATTGGATGAAAATCATAAACATTCAGTTAATA 638
QY 121 ACCGTATTTCCAGTTCAGATGAATACAGTTTGAATGAAGTGATGGCATCTAAGAAG 180
Db 639 ACAGTATTTCCAGTTCAGATGAATACAGTTTGAATGAAGTGATGGCATCTAAGAAG 698
QY 181 TTTTGAATTTTGAATCTTACATGTTGTCCTCCACTTCAGATGCTGCAGACGAA 240
Db 699 TTTTGAATTTTGAATCTTACATGTTGTCCTCCACTTCAGATGCTGCAGACGAA 758
QY 241 TTTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCCTGCAATCCAAAGCTGTGAAATT 300
Db 759 TTCTGTCCAGCAGAGAGTTGTGACAGATACGGCCTGCAATCCAAAGCGGTGAGATTG 818
QY 301 TGGCAACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAAAGAAAAGCATTTAATAA 360
Db 819 TGGCCCAAGAGATGATGACTGATTTGCCAAGCTCGTTGAAAGAAAAGCATTTAATAA 878
QY 361 TGGTGGCTTTGATATGATGAAGAGCTGCAAGAAAATGCTATGAGAAATCTGGCCTGA 420
Db 879 TGGTGGCTATGATATGATGAAGAGCTGCAAGAAAATGCTATGAGAAATCTGGCCTGA 938
QY 421 CACCAATGATATGACGTAATAGAACTTCACGATGCTTTTCTACCAACGAATCCTTA 480
Db 939 CACCAACGATGTGACGTAATAGAACTTCACGATGCTTTTCTACCAACGAATCCTTA 998
QY 481 CTTATGAAGCACTGGGACTCTGTCCAGAAGGACAAGGTGCAACGCTGTTGATAGAGAG 540
Db 999 CTTAGCAAGCACTGGGCTCTGTCCAGAAGGACAAGGTGCAACGCTGTTGATAGAGAG 1058
QY 541 ATAATCATATGAGAGAAAGTGGTCAATAATCTAGTGTGACTGATTTCAAAGGAC 600
Db 1059 ACAACACTTACGAGGAAGTGGTCAATCAACCTAGTGAAGGCTCATCTCAAGGGAC 1118
QY 601 ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAAACTGTGCTGGCAGCTGAGAGGG 660
Db 1119 ACCCACTGGGCGCAAGGCTCTGCTCAGTGGCGGAGCTCTGCTGGCAGCTGAGAGGG 1178
QY 661 AAGCCGGAAGAAGCAAGTTCCTGCTGCAAAAGTGGCTCTGCNGCATAAATTANGCAT 720
Db 1179 AAGCCGGA--GAGGCAGGTTCCCGGGCAAAAGTGGCTCTGCAGCACAATCTAGGCCT 1236

Db 1319 ATCTCATTTTAAAGAA-----ATCGAGAAGAAGCTTGAAGAGGAGGGGAAGAGTTCCG 1372

QY 906 TGAAGAAAAAATNCGNGGGAATTTTGCCTTCAAGGGGAANAATGCCCCGCGGGGT 965

Db 1373 TGAAGAAAATCGGTGC-----ATTTTGCCTTCAAGTGAAGATGCCCCGCGGGCAA 1427

QY 966 AAAAG 970

Db 1428 AGAAG 1432

RESULT 10

RAT60KDA 2599 bp mRNA linear ROD 27-APR-1993

LOCUS Rat 60 kDa protein and non-specific lipid transfer protein mRNA, complete cds.

ACCESSION M62763 GI:202552

VERSION M62763.1

KEYWORDS 60 kDa protein; non-specific lipid transfer protein.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 2599)

AUTHORS Mori,T., Tsukamoto,T., Mori,H., Tashiro,Y. and Fujiki,Y.

TITLE Molecular cloning and deduced amino acid sequence of nonspecific lipid transfer protein (sterol carrier protein 2) of rat liver: a higher molecular mass (60 kDa) protein contains the primary sequence of nonspecific lipid transfer protein as its C-terminal part

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (10), 4338-4342 (1991)

MEDLINE 91239563

PUBMED 2034675

COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley) cDNA to mRNA.

FEATURES

source location/Qualifiers

1. 2599 /organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

12. 1977 /product="clone TM-34"

46. 1689 /standard_name="clone TM-34"

/codon_start=1

/product="60 kDa protein"

/protein_id="AAA0622.1"

/db_xref="GI:202553"

/translation="MPSVALNSPRLRVFVVGVMTEKMPGGENSRDYPDLAKEAGQ KALADAQIPYSAVEQACVGVYGESTCGORAIYHSLGLGIPINNNNCSGSTALLF MAQOLVQGLANCVLALGFEMKEKSLGTYSDRSNPLEKHIDVLINKYMSACPFAP OLFSGAKEHMETVGTKEVHFAKIGMKNKHSVNNPYSOQDEYSLDEIMKSRPVDF LTVLQCCPTSDGAAAIYSSEFVQKHGLQSKAVEIYAQEMVTMPSTFEKSVIKMV GYDMSKEARKCYEKSGLGSDVDVIELHDCFSTNELITYEALGLCPGGGALVDRG DNTYGGKWINPSGGLISKHPLGATGLAQCAELCWLIRGEAGKRQVPAKYALQHNL GLGGAAVVTLYRMGFPEAASSFRTHQISAPTSAGDGFKANLIFKEIEKLEEGEE FVKKIGIFAFKVKVDGPGKEATVVDVKNKGSVLPDSDKADCTITMADSDLLALM TGMNPOSAPFQGLKTIAGNMGLAMKLSLQLQPDYKAKL"

CDS

1258. 1689 /note="larger precursor of non-specific lipid transfer carrier protein (sterol carrier protein 2)"

/codon_start=1

/product="non-specific lipid transfer protein"

/protein_id="AAA0623.1"

/db_xref="GI:202554"

/translation="MGFPEAASSFRTHQISAPTSAGDGFKANLIFKEIEKLEEGE BEFVKKIGIFAFKVKVDGPGKEATVVDVKNKGSVLPDSDKADCTITMADSDLLA LMTGMNPOSAPFQGLKTIAGNMGLAMKLSLQLQPDYKAKL"

sig_peptide 1258. 1317

mat_peptide 1318. 1686

/product="non-specific lipid transfer protein"

/note="mature non-specific lipid transfer carrier protein (sterol carrier protein 2)"

polyA_signal 1953. 1958 /note="clone TM-11 & 34"

polyA_signal 2160. 2165 /note="clone TM-28 & 33"

polyA_signal 2242. 2247 /note="clone TM-28 & 33"

polyA_signal 2382. 2387 /note="clone TM-28 & 33"

polyA_signal 2580. 2585 /note="clone TM-33"

polyA_signal 2581. 2586 /note="clone TM-33"

BASE COUNT 727 a 524 c 654 g 694 t

ORIGIN

Query Match 57.9%; Score 575.2; DB 10; Length 2599;

Best Local Similarity 78.5%; Pred. No. 3.2e-133;

Matches 758; Conservative 0; Mismatches 188; Indels 19; Gaps 6;

QY 6 CCAATTGCTCCTCAGATGTTGGGTATGCTGAAAAGAACATATGAAAAATATGAGACA 65

Db 514 CCATTGCTCCTCAGCTGTTGGAGCGCTGGGAAGAACATATGAAACATATGAGACA 573

QY 66 AAAATTGAACACTTTCGAAAATTTGGATGGAATAATCATAAACATTCAATTACCCG 125

Db 574 AAGTTGAACACTTTCGAAAATTTGGATGGAATAATCATAAACATTCAATTACCCG 633

QY 126 TATTTCCAGTTCGAAGATGATATACAGTTAGATGAAGTATGGCATCTAAGAAGTTT 185

Db 634 TATTTCCAGTTCGAAGATGATATACAGTTAGATGAAGTATGAATAAAGGCCAGTTTC 693

QY 186 GATTTTGTACTATCTTACAAATGTTGTCCCACTTCAGATGCTGCAGCAGCAATTTTG 245

Db 694 GATTTTGTACTATCTTACAAATGTTGTCCCACTTCAGATGCTGCAGCAGCAATTTTG 753

QY 246 GCCAGTGAAGCATTTGTACAGAAGTATGGCCTGCAATCCAAAGCTGTGAAATTTGGCA 305

Db 754 TCTAGTGAAGAGTTTGTGCAAGAAGCATGGCCTGCAATCCAAAGCTGTGAAATTTGGCA 813

QY 306 CAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTTAATAAATGGTT 365

Db 814 CAGGAGATGCTGACTGACATGCCCAAGTACATTTGAAGAAAAAGTGTATTAAATGGTT 873

QY 366 GCGTTGATATGAGTAAAGAAGCTGCAGAAAATGCTATGAGAAATCTGGCCTGACACCA 425

Db 874 GCGTTGATATGAGTAAAGAAGCTGCAGAAAATGCTATGAGAAATCTGGCCTGACACCA 933

QY 426 AATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGAACTCCTTACTTAT 485

Db 934 AGTATGTCGACGTGATAGACTTCACGATTGCTTTTCTAACCAAGAACTCCTGACTTAC 993

QY 486 GAAGCACTGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAGATAAT 545

Db 994 GAAGCACTGGGCTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAGACAAC 1053

QY 546 ACATATGAGAGAAAGTGGTCAATAATCCTAGTGTGAGCTGATTTCAAGGAGACACCA 605

Db 1054 ACTTACGAGAGAAAGTGGTCAATAATCCTAGTGTGAGAGCTCATCTCCAAGGAGACACCA 1113

QY 606 CTAGCGCTACAGGCTTGTCTCAGTGTGCAAGAACTCTGCTGCAGCTGAGAGGGAGGCC 665

Db 1114 CTGGGTGCCACAGGCTGTGCTCAGTGCAGGAGCTCTGCTGCAGCTGAGAGGGAGGCC 1173

QY 666 GGAAGAGAGCAAGTTCCTGCTGCAAAAGGTGGCTCTGCGNCAATAATTANGCATTTGAG 725

Db 1174 GGAAGAGAGCAAGTTCCTGCTGCAAAAGGTGGCTCTGCGNCAATAATTANGCATTTGAG 1231

QY 726 GAAGTGTGTTGTAACACTCTACAAGATGGGGTTTCCCGAAGCCGCCAGTTCCTTTTA 785

Db 1232 GAGCTGCTGTGTCAACCTCTACAAGATGGG--TTTTCGGAAGCTGCCAGCTCC--TTCA 1288

QY	786	GAATCATCAAAATTGAAAGCCNGTTCACACCAAGCTCTGCAAGTATNATNGGTTTAANGNA	845
Db	1289	GAAAGCACCAGATTTCAGCTGCTCCACACCAAGCTCTGCAAGGATGATTCAGGCCAA	1345
QY	846	AATCTNGTTTAAAGGNGATTGAGAAAGAAACNTNAAAGAGGGAANGGGGAACAATTG	905
Db	1346	ATCTCATTTTAAAGGAAATCGAGAGAAAGCTTGAAGAGAAAGGGGAAGATTG	1399
QY	906	TGAAAGAAAATNCGNGGGAATTTTTCCTTCAAGGGGAANAATGCCCCGCGGGT	965
Db	1400	TGAAGAAAATCGGTGCATTTTTCCTTCAAGTGAAGAGATGCGCCCGGGGCAA	1454
QY	966	AAAAG 970	
Db	1455	AGAAG 1459	
RESULT 11			
AX306091		2152 bp	DNA
LOCUS	AX306091	Sequence 842 from Patent WO0188188.	linear PAT 11-DEC-2001
DEFINITION	AX306091		
ACCESSION	AX306091.1	GI:17645404	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match	57.7%;	Score 573.4;	DB 6; Length 2152;
Best Local Similarity	78.9%;	Pred. No. 9.1e-133;	
Matches	765;	Conservative 0;	Mismatches 186; Indels 19; Gaps 7;
QY	1	CTCACCACAGTTGCTCCTCAGATGTTGGGTATGCTGGAAAGAACATATGAAAAATATG	60
Db	479	CACATCCGATTACTCCTCAGATGTTGGGTATGCTGGAAAGAACATATGAAAAATATG	538
QY	61	GAACAAAATTTGAACACCTTTGCAAAAATTGGATGGAATAATCATTAACATTCACTTAATA	120
Db	539	GAACAAAATTTGAACACCTTTGCAAAAATTGGATGGAATAATCATTAACATTCACTTAATA	598
QY	121	ACCCGTATTTCCAGTTCCAGATGAATACAGTTAGATGAAGTATGTCATTAAGAG	180
Db	599	ACAGTATTTCCAGTTCCAGATGAATACAGTTAGATGAAGTATGAATCAAAACAG	658
QY	181	TTTTGATTTTTGACTATCTTACAATGTTGTCCCACTTCAGATGCTGCAGCAGCAA	240
Db	659	TTTTGATTTTTGACTATCTTACAATGTTGTCCCACTTCAGATGCTGCAGCAGCAA	718
QY	241	TTTTGGCCAGTGAAGCATTTGTACAGAGTATGCGCTGCAATCCAAAGCTGTGAATTT	300
Db	719	TTCTGTCCAGCGAGAGTTGTGCAAGCAGTACGCGCTGCAATCCAAAGCGGTGAATTTG	778
QY	301	TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAAGCATTATTA	360
Db	779	TGGCCCAAGAGATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAAGCATTATTA	838
QY	361	TGGTTGGCTTTGATATGATTAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCTGA	420
Db	839	TGCTTGGCTATGATATGATTAAGAGCTGCGAGAGATGCTATGAGAAATCTGGCTGA	898

OY		421 CACCAATGATATTGAACGTAAATAGAACCTTCACGATTGGCTTTCTACTACAAGAATCCTTA	480
Db			
	899 CACCACGCATGTGCACGTGATAGAGCTTCACGATTGCTTCTGTCAATGAATCATCA	958	
OY	481 CTATGAAGCACTGGGACTCTGTCCAGAAGCAAGGTCCAAGCTGGTGATAGAGGAG	540	
Db			
	959 CTTAGAAGCACTGGGGCTCTGTCCAGAAAGCAAGGTGGAACCCGTGGTGACAGAGGGG	1018	
OY	541 ATAATACATAAGAGGAAAAGTGGGTATATAATCTAGTGTGACTGATTTCAAAGGAC	600	
Db			
	1019 ACAACACTTACGAGAGGAAAAGTGGGTATCAACCTAGTGAGGCCCTCATCTCAAAGGAC	1078	
OY	601 ACCCACTAGGCGGCTACAGGCTCTGCTCAGTGTGACAGAACTCTGCTGGCAGCTGAGAGGG	660	
Db			
	1079 ACCCACTGGGGCGCAACAGGTCTGGCTCAGTGGCGGAGACTCTGCTGGCAGCTGAGAGGG	1138	
OY	661 AAGCCGGAAAAAGGCCAAAAGTTCCTGGTGCAAGGTGGCTCTGCNGCATTAATTANGCAT	720	
Db			
	1139 AAGCCGAAA--GAGGCAGGTTCCCCGGGCCAAAGGTGGCTCTGCAGCACAAATCTAGGCC	1196	
OY	721 TGGAGAACTGTGGTGTGAACACTCTACAGATGAGGGTTTTCCCGAAGCCGCCAGTTCC	780	
Db			
	1197 CGGAGAGCTGTGGTTGTCAACCTCTACAGATGGG--TTTTCCCGAAGCTGCCAGCTCC	1254	
OY	781 TTTTAGAATCATCAAAATTGAAGCCNGTCCAACCAAGCTCTGCAAGTNATNGSTTTAA	840	
Db			
	1255 TTCAGGA--CGCACCAAGGTTTCAGCTGCTCCCA--CCAGCTCTGCAGGGGATGSAITCAA	1310	
OY	841 NGNAAAATCTNGTTTTTAAAGNGGATTGAGAGGAACNTNAAGAGGGGAAGGGGAACA	900	
Db			
	1311 GGCAAAACCTCGTCTTTAA--GGAGATTGAGAGAAGCTTGA---AGAGGAAGGGGAACA	1364	
OY	901 ATTTGTGAAGAAAATNCGGNGGAATTTTGCCTTCAAGGGGAAANAATGGCCCTGG	960	
Db			
	1365 GTTCGTGAAGAAAATCGGTGGC-----ATTTTGCCTTCAAGTGAALAGATGGCCCTGGA	1419	
OY	961 GGGGTAAAG	970	
Db	-		
	1420 GGCAAAAGAG	1429	

RESULT 12	
LOCUS	MUSSCP 2152 bp mRNA linear ROD 14-JUL-1993
DEFINITION	Mus musculus sterol-carrier protein X mRNA, complete cds.
ACCESSION	M91458
VERSION	M91458.1 GI:293793
KEYWORDS	non-specific lipid transfer protein; sterol carrier protein.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2152) Seedorf, U., Raabe, M. and Assmann, G. Cloning, expression and sequences of mouse sterol-carrier protein-X-encoding cDNAs and a related pseudogene Gene 123 (2), 165-172 (1993)
JOURNAL	93154580
MEDLINE	8428655
PUBMED	Original source text: Mus musculus (strain C57BL/ 6 x CBA) (library: 1-ZAP II) female adult, 6-8 weeks liver cDNA to mRNA.
COMMENT	Location/Qualifiers
FEATURES	1..2152
Source	/organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/ 6 x CBA" /db_xref="taxon:10090" /sex="female" /tissue_type="liver" /dev_stage="adult, 6-8 weeks" /tissue_lib="1-ZAP II" 16..1659

```
/codon_start=1
/product="sterol-carrier protein X"
/protein_id="AAA40098.1"
/db_xref="GI:293794"
/translation="MPSVALKSPRLRFRVVGVMKMKPGGENSRDYPDMAKEAQ
KLEDAQIPYSAVEQACVGYGDSTSGORAIYHSLGLGIPINNNSTGSTALF
MAHQLIQGLANCVLALGFEEKMERGSIKFSDRTPIDKHEVLIDKYGLSAHITP
OMFYAGKEHMEKYGTKEVHEPAKIGMKNKHSVNNVTSQODEYSLBVMKSPYDF
LTIIQCCPTSDGACAAIISSEFVQYGLQSKAVEIYAQEMMTDLPSTFEKSLIKVY
GYDMSKEAARCYEKGSLTPNDVDVIELHDFSVNELITYEALGLCPGOGGTLVDG
DNTYGGKVINPSGGLISKHPLGATGLAQCAELCWQLRGEAGKRQVPGAKVALQHL
GLGAVVVTLYRMGFPEAASSFRTHQVSAPTSAGDGRKANLVEKEIEKLEEGEQ
FVKKIGGIFAFKVKDGPFGKEATWVDVKNKGSVLPNSDKKADCTTMADSDLLALM
TGKMPQSAFFQGLKTIAGNMGLAMKQLQLQPGKAKL"
```

polyA_signal 2137..2142
polyA_site 2152
BASE COUNT 586 a 455 c 576 g 535 t
ORIGIN

```
Query Match 57.7%; Score 573.4; DB 10; Length 2152;
Best Local Similarity 78.9%; Pred. No. 9.1e-133;
Matches 765; Conservative 0; Mismatches 186; Indels 19; Gaps 7;

Qy 1 CTCACCCAGTTGCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 60
Db 479 CACATCCGATTACTCCTCAGATGTTGGGTATGCTGGAAAAAGACATATGAAAAATATG 538

Qy 61 GAACAAAAATGAACACTTTGCAAAAATTCGATGAAAAATCATAAACATTCACTTAATA 120
Db 539 GAACAAAAATGAACACTTTGCAAAAATTCGATGAAAAATCATAAACACTCACTTAATA 598

Qy 121 ACCCGATTCCAGTTCCAGATGAATACAGTTTAGATGAAGTGAATGCGATCTAAGAAG 180
Db 599 ACACGTAATCCAGTTCACAGATGAATACAGCTTAGAGAAGTAATGAATAACAAACCAG 658

Qy 181 TTTTGTATTTTGTACTATCTTACATGTTGTTCCACTTCAGATGCTGCTGCAGCAGCAA 240
Db 659 TTTTCGATTTTCTGACTATCTTCAATGCTGCCACCCTCAGATGCTGCTGCCGCCGA 718

Qy 241 TTTTGGCAGTGAAGCATTGTACAGAGTATGGCCCTGCAATCCAAAGCTGTGAAATTT 300
Db 719 TTCTGTCCAGCAGAGAGTTGTGACAGATACGGCCCTGACATCCAAAGCGGTGAGATTG 778

Qy 301 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAAAA 360
Db 779 TGGCCCAAGAGATGATGACTGACTTACCAGTACGTTTGAAGAAAAAGTATTATTAAAG 838

Qy 361 TGGTGGCTTGTATGAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA 420
Db 839 TGGTGGCTATGATATGAGTAAGAAGCTGCCAAGAGATGCTATGAGAAGTCCGGCCTGA 898

Qy 421 CACCAATGATATTTGACGTAATAGAACTTCAAGATGCTTTTCTAACCAACGAACCTCTTA 480
Db 899 CACCCAACGATGTCGACGTGATAGAGCTTCAAGATTGCTTCTGTCAATGAACCTCATCA 958

Qy 481 CTTATGAAGCACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 540
Db 959 CTTACGAAGCACTGGGGCTGTGTCCAGAAAGCAAGGTGGAACCTGTGTGACAGAGGGG 1018

Qy 541 ATAATACATATGAGGAAAGTGGGTCAATAAATCTAGTGTGAGCTGATTTCAAGGGAC 600
Db 1019 ACAACACTTACGAGGAAAGTGGGTCAATCAACCTAGTGAGGCTCATCTCAAGGGAC 1078

Qy 601 ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGACAGAACTCTGCTGGCAGCTGAGAGGG 660
Db 1079 ACCCACTGGGCGCAACAGGCTGTGGCTCAGTGGCGGAGCTCTGCTGGCAGCTGAGAGGG 1138

Qy 661 AAGCCGGAAGAAAGCAAGTTCCTGGTGCAAGGTGGCTCTGCNGCATATTTANGCAT 720
Db 1139 AAGCCGGAAGAA--GAGGCAAGTTCCTGGGCGCAAGGTGGCTCTGCAAGCAATCTAGGCC 1196

Qy 721 TGGAGGAATGTGTTGTAACACTCTACAGATGGGGTTTCCGGGAAGCCGCGCAGTTCC 780
```

```
Db 1197 CGGAGAGCTGTGTTGTACACCTCTACAGATGGG--TTTCCCGAAGCTGCAGCTCC 1254
Qy 781 TTTAGACTCATCAAAATGAAAGCCNGTTCACCAAGCTCTGCAAGTATNGGTTTAA 840
Db 1255 TTCAGGA--CGACACAGGTTTCAGCTGCTCCCA--CGAGCTCTGCAGGGATGATTCAA 1310
Qy 841 NGNAAATCTNGTTTAAAGNGGATTTGAGAAGGAACNTNAAAGAGGGAANGGGACA 900
Db 1311 GGCAAACCTGCTCTTAA--GAGATTTGAGAAGAGAGCTTGA---AGAGGAAGGGGACA 1364
Qy 901 ATTGTGAAAGAAAAATNCGNGGGAATTTTGCCCTTCAGGGGGAANAATGGCCCTGG 960
Db 1365 GTTCGTGAAGAAATCGGTGC---ATTGTCCTTCAAGTGAAGATGGCCCTTGA 1419
Qy 961 GGGGTAAAG 970
Db 1420 GCAAGAAAG 1429
```

```
RESULT 13
RATSCP2 1409 bp mRNA linear ROD 27-APR-1993
LOCUS Rat sterol carrier protein-2 (SCP-2) mRNA, complete cds.
DEFINITION M34728
ACCESSION M34728.1 GI:206871
VERSION M34728.1 GI:206871
KEYWORDS sterol carrier protein-2.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1409)
AUTHORS Billheimer,J.T., Strehl,L.L., Davis,G.L., Strauss,J.F. III and
Davis,L.G.
TITLE Characterization of a cDNA encoding rat sterol carrier protein2
JOURNAL DNA Cell Biol. 9 (3), 159-165 (1990)
MEDLINE 90253610
PUBMED 2340090
COMMENT Original source text: Rat liver, cDNA to mRNA, clone SP43.
FEATURES
source
1..1409
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
<1..1409
/product="sterol carrier protein-2 mRNA"
307..1128
/note="sterol carrier protein-2 (SCP-2)"
/codon_start=1
/protein_id="AAA42120.1"
/db_xref="GI:206872"
/translation="MWGYDMSKEAARKCYEKSGLPSDVVIELHDFSNELLTYEA
LGLCPGEGGALVDRGNTYGGKVINPSGLISKHPLGATGLAQCAELCWQLRGEA
GKRQVPGAKVALQHLNGLGGAAVVTLYRMGFPEAASSFRTHQISAAPRSSAGDFKAN
LIFKEIEKLEEGEEFVKKIGIFAFKVKDGPFGKEATWVDVKNKGSVLPNSDKK
ADCTITMADSDLLALMTGKMPQSAFFQGLKTIAGNMGLAMKQLQLQPDKAKL"
```

```
BASE COUNT 387 a 300 c 390 g 332 t
ORIGIN 2 bp upstream of EcoRI site.

Query Match 52.5%; Score 522.2; DB 10; Length 1409;
Best Local Similarity 77.8%; Pred. No. 6.1e-120;
Matches 713; Conservative 0; Mismatches 183; Indels 20; Gaps 7;
```

```
Qy 55 AATATGAACAAAAATTTGAACACTTTGCAAAAATTCGATGAAAAATCATAAACATTGAG 114
Db 3 AATTCGAACAAAGTTGAACACTTTGCAAAAATTCGATGAAAAATCATAAACACTGAG 62
Qy 115 TTAATAACCCGATTTCCAGTTCACAGATGAATACAGTTTAGATGAAGTGAATGATCTA 174
Db 63 TTAATAACCCGATTTCCAGTTCACAGATGAATACAGCTTAGATGAGATGAATGAATCAA 122
Qy 175 AAGAAGTTTGTGATTTTGACTATCTTACAAATGTTGTCCTCACTTCAGATGGTGTGAG 234
```

Db 123 GGCCAGTTTTCGATTTTCTGACTGTCTTACAATGCTGTCCCACTCAGATGTGCCGAG 182

QY 235 CAGCAATTTTGGCCAGTGAAGCAATTGTACAGAAATATGCGCTGCAATCCAAAGCTGTG 294

Db 183 CAGCAATTTGTGTCTAGTAGAGATTGTGCAGAAAGCATGCGCTGCAGTCCAAAGCTGTG 242

QY 295 AAATTTTGGCACAGAATAATGATGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTA 354

Db 243 AAATTGTGGCACAGAGATGTGTGACATGCCCCAGTACATTTGAAG-AAAAAGTTTA 301

QY 355 TTAAATGGTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTG 414

Db 302 TTAAATGGTTGGCTATGATATGATTAAGAAAGCTGCCAGAAAGTCTATGAGAAAGTCCG 361

QY 415 GCCTGACACCAAAATGATATTGACGTAAAGAACTTACAGATTGCTTTTCTACCAACGAA 474

Db 362 GCGTGGGTCCAGTGTATGACGTGATAGAGCTTACAGATTGCTTCTTACCAATGAC 421

QY 475 TCCTTACTTATGAGCACTGGGACTGTCTCAGAAAGACAAAGTGCAACGCTGTGATTA 534

Db 422 TCCTGACTTATGAGCACTGGGCTCTGTCCAGAAAGACAAAGTGAGCACTGTGAGCA 481

QY 535 GAGGAGATATATACATATGAGAGAAAGTGGGTCTAATAATCCTAGTGTGACTGATTTCA 594

Db 482 GAGGGACACAACACTTACGAGAGAAAGTGGGTCTAATAACCTAGTGAAGGCTCATCTCCA 541

QY 595 AGGACACCCCACTAGGCGCTACAGGTCTGTCTAGTGTGCAAACTCTGCTGGCAGCTGA 654

Db 542 AGGACACCCCACTAGGCTGCTGACAGGTCTGCTCAGTGCAGGAGCTCTGCTGGCAGCTGA 601

QY 655 GAGGGAGAGCCGGAAAAAGAGCAAAAGTCTCTGTGCAAAAGTGGCTCTGCNGCATTAATT 714

Db 602 GAGGCGAAAGCCGAAA--GAGGCAAGTCTCTGGGCAAAAGTGGCTCTGCAGCACAATTT 659

QY 715 ANGCAATTGAGGAAGTGTGTGTTAACACTCTACAAGATGGGTTTCCCGAAGCCGCC 774

Db 660 AGGCTTGGAGAGAGTGTGTGTTCACCTCTACAGAAATGGG--TTTCCCGAAGCTGCC 717

QY 775 AGTCTTTTGAAGTCACTCAAAATTGAAGCCNGTCCAAACCAAGCTCTGCAAGTNAATNG 834

Db 718 AGCTCC-TTCAGAACGACCAAGATTTC--GCTGCTCCCAACAGCTCTGCAGGGGATGG 773

QY 835 GTTAAANGNAAAATCTNGTTTAAAGNGGATTGAGAGAAACNTNAAAAGAGGAANG 894

Db 774 ATTCAAGGCAAAATCTCATTTTAAAGAA-----ATCGAAGAAAGCTTGAAGAGGAAG 827

QY 895 GGAACAATTTGTGAAGAAAAAATNCGNGGGAATTTTGCCCTTCAAGGGGAAAAAATGG 954

Db 828 GGAAGATTCTGTAAGAAAAATCGGTGC-----ATTTTGCCTTCAAAAGTGAAGATGCG 882

QY 955 CCCTGGGGGTAAAG 970

Db 883 CCGGGGGCAAGAAG 898

RESULT 14

HUMSCP2B 1500 bp mRNA linear PRI 06-DEC-1993

LOCUS HUMSCP2B 1500 bp mRNA linear PRI 06-DEC-1993

DEFINITION Human sterol carrier protein 2 mRNA, complete cds.

ACCESSION M75884

VERSION M75884.1 GI:432976

KEYWORDS sterol carrier protein-2.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1500)

AUTHORS He,Z., Yamamoto,R., Furch,E.E., Schantz,L.J., Naylor,S.L., George,H., Billheimer,J.T. and Straus,J.F. III.

TITLE cDNAs encoding members of a family of proteins related to human sterol carrier protein 2 and assignment of the gene to human chromosome 1 p21---pter

JOURNAL DNA Cell Biol. 10 (8), 559-569 (1991)

MEDLINE 92029618

PUBMED 1718316

REFERENCE 2 (bases 1 to 1500)

AUTHORS Vesa,J., Hellsten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T., Mead,S., Cowell,J.K., Straus,J.F.III, and Peltonen,L.

TITLE Assignment of sterol carrier protein X/sterol carrier protein 2 to 1p32 and exclusion as the causative gene for infantile neuronal ceroid lipofusiosis

JOURNAL Unpublished

COMMENT On Dec 6, 1993 this sequence version replaced gi:337996.

ORIGINAL source text: Human liver cDNA to mRNA.

FEATURES

location/Qualifiers

1. .1500

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/map="1p32"

/issue_type="liver"

/clone_1fb="Clontech HL-1001b,HL-1115b"

1. .1500

/gene="SCP-2"

82. .951

/gene="SCP-2"

/codon_start=1

/product="sterol carrier protein-2"

/protein_id="AAA03558.1"

/db_xref="GI:432977"

/translation="MTDLPSFEEKSIIKMGFDMSKEARKCYEKSGLTPNDIDVI ELHDCFSTNELITYEALGLCPGOGATLVDRDNTYGGKVINPSGGLISKHPLGAT GLAQCAELCWOLRGEAKROVPGAKVALQHNLGIGAVVTLTKMGFPEAASSFRTHQ IEAVPTSSASDFKANLVFKEIEKLEEEQFVKIKIGIFAFKVKDGGKEATWV DVKNGKGSVLPNSDKADCTITMADSDFLALMTGKNPQSAFPGKLKITGMGLAMK LQNLQLPGNAKL"

1218. .1223

polyA_signal

BASE COUNT 482 a 254 c 342 g 422 t

ORIGIN

Query Match 51.1%; Score 508.2; DB 9; Length 1500;

Best Local Similarity 87.8%; Pred. No. 2e-116;

Matches 640; Conservative 0; Mismatches 69; Indels 20; Gaps 8;

QY 257 ATTGTACAGAAATATGCGCTGCAATCCAAAGCTGTGAAATTTTGGCACAGAAATGAT 316

Db 27 ATTTTACAGAAATATGCGCTGCAATCCAAAGCTGTGAAATTTTGGCACAGAAATGAT 86

QY 317 GACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTAATAATGTTGGCTTTGATAT 376

Db 87 GACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTAATAATGTTGGCTTTGATAT 146

QY 377 GAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCGCCCTGACACCAATGATATGA 436

Db 147 GAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCGCCCTGACACCAATGATATGA 206

QY 437 CGTAATGAACCTCAGATTGCTTTTCTACCAACGAACTCCTTACTTATGAAGCACTGG 496

Db 207 CGTAATGAACCTCAGATTGCTTTTCTACCAACGAACTCCTTACTTATGAAGCACTGG 266

QY 497 ACTCTGTCAGAAAGGACAAAGGTGCAACGCTGTTGATAGAGAGATATACATATGAGG 556

Db 267 ACTCTGTCAGAAAGGACAAAGGTGCAACGCTGTTGATAGAGAGATATACATATGAGG 326

QY 557 AAAGTGGTCATAAATCTAGTGTGAGTGAATTTCAAAGGACACCCACTAGGCGCTAC 616

Db 327 AAAGTGGTCATAAATCTAGTGTGAGTGAATTTCAAAGGACACCCACTAGGCGCTAC 386

QY 617 AGTCTTGTCTCAGTGTGCAAACTCTGTGCGCAGCTGAGAGGGAAGCCGAAAAAGAGC 676

Db 387 AGTCTTGTCTCAGTGTGCAAACTCTGTGCGCAGCTGAGAGGGAAGCCGAAA--GAGG 444

OY		677	AAAGTTCCTGGTGCAAGGTGCTCTGCNGCATATTATTCAGACTTGAGGAACGTGGTT	736
Db		445	CAAATCCCTGGTGCAAGGTGCTCTGCAGCATATTATTCAGACTTGAGGAGCTGGTT	504
OY		737	GTAACACTCTACAAGATGGGTTTTCCCGAAGCCGCCAGTTCCTTTAGAACATCAAA	796
Db		505	GTAACACTCTACAAGATGGG--TTTCCGGAAGCCGCCAGTT-CTTTAGAACTCATCAA	561
OY		797	AATTGAAGCNGTTCACCAAGCTCTGCAAGTNATNGTTTAAANGNAAAATCTNGTTT	856
Db		562	ATTGAA--GCTGTTCCAACACAGCTCTGCAAGTAGTAGTTTAAGCAATCTGTGTTT	618
OY		857	AAAGNGGATTGAGAAGGAAACNTNAAAGAGGGAANGGGGAACAATTGTGAAGAAAAA	916
Db		619	AA--CGAGATTGAGAAGAACTTGA---AGAGGAAGGGGAACAAGTTGTGAAGAAAAATC	672
OY		917	TNCGNGGGAATTTTGGCCCTTCAAGGGGAAANAATGCCCTGGGGGTTAAAGANGGCC	976
Db		673	GGTGGT----ATTGTTGCTTCAAGGTGAAG-ATGGCCTGGGGGTAAAGAGGCCACC	726
OY		977	ACCCTGGGG 985	
Db		727	TGGGTGGTG 735	
RESULT 15				
CHKSCP2A				
LOCUS	CHKSCP2A	2443 bp	mRNA	linear VRT 11-AUG-1993
DEFINITION	Gallus gallus domesticus sterol carrier protein-2 mRNA, 3' end.			
ACCESSION	L09231			
VERSION	L09231.1 GI:304422			
KEYWORDS	sterol carrier protein-2.			
SOURCE	Gallus gallus (chicken)			
ORGANISM	Gallus gallus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
	Phasianinae; Gallus.			
	1 (bases 1 to 2443)			
REFERENCE	Pfeifer,S.M., Sakuragi,N., Ryan,A., Johnson,A.L., Deeley,R.G.,			
AUTHORS	Bilhheimer,J.T., Baker,M.E. and Strauss,J.F. III.			
TITLE	Chicken sterol carrier protein 2/sterol carrier protein x: cDNA cloning reveals evolutionary conservation of structure and regulated expression			
JOURNAL	Arch. Biochem. Biophys. 304 (1), 287-293 (1993)			
MEDLINE	93312016			
PUBMED	8323294			
COMMENT	Original source text: Gallus gallus (strain White Leghorn, sub_species domesticus) (library: lambda gt11) adult liver cDNA to mRNA.			
FEATURES				
source	Location/Qualifiers			
	1..2443			
	/organism="Gallus gallus"			
	/mol_type="mRNA"			
	/strain="White Leghorn"			
	/sub_species="domesticus"			
	/db_xref="taxon:9031"			
	/tissue_type="liver"			
	/dev_stage="adult"			
	/tissue_lib="lambda gt11"			
	<1..1646			
	/standard_name="SCP2"			
	/codon_start=3			
	/product="sterol carrier protein-2"			
	/protein_id="AAA02488.1"			
	/db_xref="GI:304423"			
	/translation="ARVCEERGGPRAIMQRRVVVGVMTKFAKPSSENSVDYPDLAKE			
	AGOKALADAGIPYSAVEOACVGVYGDSCTGORAIVHGLGTGPIINVNNCATGST			
	ALFMSRQLVEGLADCVLALGFERMAKSGLSGFSDRTPMDKHLEIMNKYGLASAP			
	ITPOMFANAGKEHMEKYGTNPYFAKIAWKNSHSTNNPYSOFOKKYTLDDEVLOSRYV			
	FDFLTVLQQCPSTNSGAAAAIILASEDFVKRHKLQPAVEILAQVMATDYPSTFEENSCM			
	KMGVDMTKKAACEKCFKAKLKPTDVDVTEIHDCFSVNEFITYEALGLCPBGACDLI			
	DRGDNITYGGKWVNPSSGLISKGHPLGATGLAQSAELCWOLRGLAGREVGGARRALO			

Query Match	42.8%	Score 425.6	DB 5	Length 2443
Best Local Similarity	73.3%	Pred. No. 9.6e-96		
Matches 557	Conservative 0	Mismatches 201	Indels 2	Gaps 1
BASE COUNT	715 a	463 c	608 g	657 t
ORIGIN	HNLGLGAAVVTLYAMGPFGAASDGVTAVPLSAVDGFKSHLVFEKIEKLOEBEEO FVKKIGVFAFKIKDGPBGKEATWVDYKNGKSVAVNSDKKADCTITMADTDLALM TGKMPOTAFPFQGLKXISGNMGMMXKLQLOPGKAKL"			
QY	6	CCAGTTGCTCCTCAGATGTTGGGTATGCTGGAAAAAGAACATATGGAATAATGGAACA	65	
Db	480	CCAATAACACCGCAGATGTTGCCAATGCTGGCAAAAGAACATATGAGAAATACGGGACA	539	
QY	66	AAAATTGAACACTTTGCCAAAAATTGGATGGAATAATCATAAACATTCACTTAATAACCCG	125	
Db	540	AATCCAGAGTACTTTGCCAAAAATTGCATGGAAGAAATCACAGCCATTCAACCAACAAACCC	599	
QY	126	TATTCACAGTTCCAGATGAATACAGTTTGAATGAAGTGAATGATGGCATCTAAGAGTTT	185	
Db	600	TATCCAGTTCCAGAGAATAACATACATTAAGTGAAGTCTCGAGTCTCGCAAGTTT	659	
QY	186	GATTTTGTGACTAATCTTACAATGTTGTCCCACTTCAGATGTTGCTGCAGACGAATTTTG	245	
Db	660	GATTTCTGACTGCTCTTACAGTGTGTCTCCACATCGAATGGCCGCCGACGTGCAATTTTG	719	
QY	246	GCCAGTGAAGCATTGTGTACAGAAGTATGGCTGCAATCCAAAGCTGTGGAATTTTGCA	305	
Db	720	GCTAGTGAAGACTTTGTGAAAAAGGCATTAAGCTTACAGCCTCAAGCTGTGGAATTCCTGGCC	779	
QY	306	CAAGAAATGATGACTGATTTGGCCAAAGCTCGTTTGAAGAAAAAAGCATTATTAATGTT	365	
Db	780	CAGGTATGGCTACTGATTATCCCAAGCAGTTTGAAGAGAACAGTTGATGAAGATGTT	839	
QY	366	GGCTTTGATATGAGTAAAGAGCTGCAGAAATAATGCTATGAGAAATCTGGCTGACACCA	425	
Db	840	GGTATGATATGACGAAAAAGCTGCAGAAAAATGCTTAAAAAAGCAGGCTAAAAACCT	899	
QY	426	AATGATATGACGTAATAGAACTTCACGATTCCTTTCTACCAACGAACCTCTTACTTAT	485	
Db	900	ACTGATGTCGATGTGATTTGAACCTCCATGACTCTTCTCAGTTAATGAGTTCAATTAACCTAC	959	
QY	486	GAAACACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAGATAAT	545	
Db	960	GAACTCTTGGACTCTGCCAGAAAGGAAAGCGTGTGACCTGATTGACAGAGAGACAAT	1019	
QY	546	ACATATGAGGAAAGTGGTCAATAATCTTATGTTGAGTGAATTTCAAAAGGACACCCA	605	
Db	1020	ACCTACGGGGGAAAGTGGTCAATAATCTTACGGGCGCTGATTTCAAAAGGACACCCCT	1079	
QY	606	CTAGCGCTACAGGTTCTTGTCTCAGTGTGCAAACTCTGCTGTCAGCTGAGAGGGAAAGCC	665	
Db	1080	CTTGTGTCACAGGCTTGGCGCAGTCCGCTGAACCTGCTGGCAGCTGCGGGCCTGGCC	1139	
QY	666	GGAAGAGAGGCAAGTTCCTGTGTGCAAAAGTGGCTCTGCNGCATATTTANGCATTTGAG	725	
Db	1140	GG--GAGGCGGGAAGTGGCGGCGCACCGCGGGCGCTGACAGCACAACTGGGCTCGGCG	1197	
QY	726	GAACTGTGTTGTAACACTTACAAAGATGGGTTTCCCG	765	
Db	1198	GGGCTGTGTGTGAGCGCTGTACGCCATGCGCTTCCCGG	1237	

Search completed: November 27, 2003, 10:02:39
Job time : 4031.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:55:53 ; Search time 2679.01 Seconds
(without alignments)
9017.749 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 994

Sequence: 1 ctcaccacgtgtgtctccacg.....ccaccctggggtgtggtggat 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_htc:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_htc:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rnd:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	688.2	69.2	817	9	AU121360 AU121360
2	655.4	65.9	828	14	CD104768 AGENCOURT
3	607.4	61.1	609	14	CB147601 K-EST0203
4	588	59.2	600	10	BE887512 601508161

5	576.6	58.0	1969	11	BC029332	BC029332 Homo sapi
6	576.6	58.0	1974	11	AK004860	AK004860 Mus muscu
7	572.6	57.6	926	10	BG289834	BG289834 602384738
8	557.4	56.1	943	10	BG105684	BG105684 602312282
9	532.4	53.6	668	9	AV653303	AV653303 AV653303
10	514	51.7	757	14	CB951253	CB951253 AGENCOURT
11	512	51.5	866	14	CB950754	CB950754 AGENCOURT
12	508.8	51.2	914	9	AI525668	AI525668 PT1.3_04
13	501.4	50.4	583	9	AL694422	AL694422 DKFZp313G
14	495.4	49.8	577	9	AL694509	AL694509 DKFZp313G
15	470.2	47.3	1051	12	BM552173	BM552173 AGENCOURT
16	469.2	47.2	525	10	BE971366	BE971366 601651537
17	457.2	46.0	763	14	CB316398	CB316398 AGENCOURT
18	443.4	44.6	599	14	CB162403	CB162403 K-EST0222
19	439.6	44.2	625	14	CB153363	CB153363 K-EST0210
20	430.8	43.3	731	14	CB947142	CB947142 AGENCOURT
21	419.6	42.2	849	10	BG432823	BG432823 602496031
22	407	40.9	1039	9	AL551168	AL551168 AL551168
23	404	40.6	532	9	AI116787	AI116787 ue21a04.y
24	386.6	38.9	545	14	R84426	R84426 yq22c09.r1
25	370.2	37.2	921	13	BX433336	BX433336 BX433336
26	368.4	37.1	805	13	BU962149	BU962149 AGENCOURT
27	362.2	36.4	1475	10	BF180638	BF180638 601808747
28	358.6	36.1	526	14	CB158882	CB158882 K-EST0218
29	356.6	35.9	744	14	CB948021	CB948021 AGENCOURT
30	355.2	35.7	456	14	F14816	F14816 SSC20E06 PO
31	349.2	35.1	532	9	AA105592	AA105592 mm65f11.r
32	341	34.3	669	12	BI147497	BI147497 602914096
33	332.8	33.5	878	10	BF699675	BF699675 602127088
34	324.4	32.6	964	10	BG289606	BG289606 602381573
35	323.6	32.6	678	10	BE911641	BE911641 601663325
36	321.6	32.4	757	10	BG718743	BG718743 602696877
37	305.4	30.7	764	14	CB947576	CB947576 AGENCOURT
38	304.4	30.6	426	14	CB286997	CB286997 CMD50_G02
39	303.4	30.5	461	14	CB286996	CB286996 CMD50_G01
40	302.6	30.4	321	9	AA343982	AA343982 EST49908
41	296.4	29.8	721	14	CD241357	CD241357 AGENCOURT
42	294.4	29.6	656	12	BI145089	BI145089 602909166
43	293	29.5	908	13	BQ877242	BQ877242 AGENCOURT
44	290.6	29.2	805	12	BG777467	BG777467 602664789
45	279.4	28.1	410	9	AW511640	AW511640 xu80a08.x

ALIGNMENTS

RESULT 1	AU121360	817 bp	mRNA	linear	EST 01-AUG-2002
LOCUS	AU121360				
DEFINITION	AU121360 HEMBB1 Homo sapiens CDNA clone HEMBB1002667 5', mRNA				
ACCESSION	AU121360				
VERSION	AU121360.1	GI:10936595			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
	1 (bases 1 to 817)				
	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,				
	Yanamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and				
	Isogai,T.				
TITLE	HRI human CDNA project				
JOURNAL	Unpublished				
COMMENT	Contact: Takao Isogai				
	Genomics Laboratory				
	Helix Research Institute				
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan				
	Tel: 81-438-52-3975				
	Fax: 81-438-52-3986				
	Email: genomics@hri.co.jp				
	HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix				
	Research Institute; CDNA library construction; Department of				

Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
source

1. .817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEMBB1002667"
/tissue_type="whole embryo, mainly body"
/dev_stage="embryo, 10 weeks"
/clone_lib="HEMBB1"
/note="Vector: pME18SFL3"

BASE COUNT 255 a 141 c 201 g 217 t 3 others
ORIGIN

Query Match 69.2%; Score 688.2; DB 9; Length 817;
Best Local Similarity 93.4%; Pred. No. 2.8e-131;
Matches 756; Conservative 0; Mismatches 45; Indels 8; Gaps 4;

QY 75 CACTTGGCAAAATTTGGATGGAAAAATCATAAACATTGATTAATACCCGATTTCCAG 134
Db 1 CACTTGGCAAAATTTGGATGGAAAAATCATAAACATTGATTAATACCCGATTTCCAG 60
QY 135 TTCCAAGATGAATACAGTTAGATGAGTGGCATCTAAAGAGTTTGTATTTTGG 194
Db 61 TTCCAAGATGAGTACAGTTAGATGAGTGGCATCTAAAGAGTTTGTATTTTGG 120
QY 195 ACTATCTTACAATGTTGTCCTCCACTTCAGATGCTGTCAGACGCAATTTGGCCAGTGA 254
Db 121 ACTATCTTACAATGTTGTCCTCCACTTCAGATGCTGTCAGACGCAATTTGGCCAGTGA 180
QY 255 GCATTGTGACAGAAGTATGCGCTGCAATCCAAGCTGTGAAATTTTGGCACAAGAAATG 314
Db 181 GCATTGTGACAGAAGTATGCGCTGCAATCCAAGCTGTGAAATTTTGGCACAAGAAATG 240
QY 315 ATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTTATTAATTTGGTGGCTTGTAT 374
Db 241 ATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTTATTAATTTGGTGGCTTGTAT 300
QY 375 ATGAGTAAAGAAAGCTGCAAGAAAAATGCTATGAGAAATCTGGCTGACACCAATGATATT 434
Db 301 ATGAGTAAAGAAAGCTGCAAGAAAAATGCTATGAGAAATCTGGCTGACACCAATGATATT 360
QY 435 GACGTAATAGAACTTCACGATTCCTTTCTACCAACGAACCTCTTACTTATGAAGCACTG 494
Db 361 GACGTAATAGAACTTCACGATTCCTTTCTACCAACGAACCTCTTACTTATGAAGCACTG 420
QY 495 GGACTCTGTCCAGAAGGACAAAGCTGCTGTTGATAGAGAGATATACATATGGA 554
Db 421 GGACTCTGTCCAGAAGGACAAAGCTGCTGTTGATAGAGAGATATACATATGGA 480
QY 555 GGAAGTGGGTCAATAATCCTAGTGTGAGCTGATTTCAAGGGGACACCCACTAGGCGCT 614
Db 481 GGAAGTGGGTCAATAATCCTAGTGTGAGCTGATTTCAAGGGGACACCCACTAGGCGCT 540
QY 615 ACAGGTCCTGCTCAGTGTGAGAACTCTGCTGGCAGCTGAGAGGGGAAAGCCGGAAGAG 674
Db 541 ACAGGTCCTGCTCAGTGTGAGAACTCTGCTGGCAGCTGAGAGGGGAAAGCCGGAAGAG 598
QY 675 GCAAAGTCTCTGGTGCAAGGTGGCTCTGCNGCATTAATTNANGCATTTGAGAGAACTGTGG 734
Db 599 GCAAAGTCTCTGGTGCAAGGTGGCTCTGCAGCATTAATTNANGCATTTGAGAGAGCTGTGG 658
QY 735 TTGTAACACTCTACAGATGGGTTTCCCGGAAGCCGCAAGTCTTTTAGAATCTCATC 794
Db 659 TTGTAACACTCTACAGATGGG--TTTTCGGAAGCCGCGCAAGTCTTTTAGAATCTCATC 716
QY 795 AAAAATGAAGCCNGTTCACAACCAAGCTCTGCAAGTNATNGTTTAANGNAAA--TCTN 851
Db 717 -AAATTGGAACCTGNTCCAAACCCAGCTTTTGCAGTGATGATTTAAGGCNAAATCTTTG 775
QY 852 GTTTTAAAGNGGATTGAGAGAAACNT 880

Db 776 GTTTTAAAGAGAAATTGAGAAAGAAACT 804

RESULT 2
CD104768 828 bp mRNA linear EST 15-MAY-2003
LOCUS
DEFINITION
CD104768
AGENCOURT_14007012 NIH_MGC_186 Homo sapiens cDNA clone
IMAGE:30372857 5', mRNA sequence.
ACCESSION
CD104768
CD104768.1 GI:30757942
VERSION
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaphs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDCM165 row: 1 column: 18
High quality sequence stop: 463.

FEATURES

source

1. .828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30372857"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_186"
/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI
(ggccattacggcc); Site 2: SfiI (ggcgcctcgcc); Library is
oligo-dT primed and directionally cloned. cDNA was
prepared from a pooled samples of tissues from skin,
meninges, duramater, pia matter and choroid plexus. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-3' and 3'
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-3' (30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library"

BASE COUNT 261 a 156 c 195 g 216 t
ORIGIN

Query Match 65.9%; Score 655.4; DB 14; Length 828;
Best Local Similarity 94.9%; Pred. No. 1.5e-124;
Matches 719; Conservative 0; Mismatches 34; Indels 5; Gaps 4;

QY 60 GGAACAAAAATTTGAACACTTTGCAAAAAATTTGATGAAAAATCATAAACATTGATTAAT 119
Db 3 GGAACAAAAATTTGAACACTTTGCAAAAAATTTGATGAAAAATCATAAACATTGATTAAT 62
QY 120 AACCCGATATCCAGTCCAGATGAATGATGATGATGATGATGATGATGATGATGATGAT 179
Db 63 AACCCGATATCCAGTCCAGTCCAGATGAATGATGATGATGATGATGATGATGATGATGAT 122
QY 180 GTTTTGTATTTTGAATCTTACATGTTGTCCCACTTCAGATGCTGTGACAGCA 239
Db 123 GTTTTGTATTTTGAATCTTACATGTTGTCCCACTTCAGATGCTGTGACAGCA 182
QY 240 ATTTGGCCAGTGAAGCATTTGTACAGAAGTATGCGCTGCAATCCAAGCTGTGAAAT 299

Db 183 ATTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCTTGCAATCCAAAGCTGTGAAATT 242

Qy 300 TTGGCACAAGAATGATGACTGATTTGGCCAAGCTCGTTTGAAGAAAAAGCATTATTAA 359

Db 243 TTGGCACAAGAATGATGACTGATTTGGCCAAGCTCGTTTGAAGAAAAAGCATTATTAA 302

Qy 360 ATGCTGGCTTTGATATGAGTAAGAAGCTGCAAGAAAAATGCTATGAGAATCTGGCTG 419

Db 303 ATGCTGGCTTTGATATGAGTAAGAAGCTGCAAGAAAAATGCTATGAGAATCTGGCTG 362

Qy 420 ACACCAATGATATTGACCGTAATAGAACTTCACGATTGCTTTCTACCAACGAATCTCTT 479

Db 363 ACACCAATGATATTGACCGTAATAGAACTTCACGATTGCTTTCTACCAACGAATCTCTT 422

Qy 480 ACTTATGAAGCACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTGATAGAGA 539

Db 423 ACTTATGAAGCACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTGATAGAGA 482

Qy 540 GATTAATACATATGAGAGAAAGTGGTCAATAATCCTAGTGGTGAAGTATTTCAAAGGA 599

Db 483 GATTAATACATATGAGAGAGTGGTCAATAATCCTAGTGGTGAAGTATTTCAAAGGA 542

Qy 600 CACCCACTAGGCGCTACAGGCTTCTGCTCAAGTGTGCAAGCTCTGCGCAGCTGAGAGG 659

Db 543 CACCCACTAGGCGCTACAGGCTTCTGCTCAAGTGTGCAAGCTCTGCGCAGCTGAGAGG 602

Qy 660 GAAGCCGGAAGAGAGCAAGTCTGCTGCTGCAAGGTGG-CTCTGCGCATTAATTANCG 718

Db 603 GAAGCCGGAAGAGAGCAAGTCTGCTGCTGCAAGGTGGCCCTGCAAGCATTAATTNAGC 662

Qy 719 ATTGAGGAA--CTGTGTTGTACA-CTCTACAGATGGGGTTTCCCGG-AAAGCGCC 774

Db 663 ATTGAGGAGAGCTGTGTTGTACACCTTAACAAATGGGGTTTCCCGAAGCCGCC 722

Qy 775 AGTTCCTTTTAGAAGTCAATAATGAAGCCNGTTC 812

Db 723 AGTTCCTTTTAGAAGTCAATAATGAAGCTGTTT 760

RESULT 3

CB147601 609 bp mRNA linear EST 29-JAN-2003

LOCUS K-EST0203627 L11SNU35481 Homo sapiens cDNA clone L11SNU35481-29-G02

DEFINITION 5', mRNA sequence.

ACCESSION CB147601 GI:28129068

VERSION CB147601

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 609)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontler Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 29 row: G column: 02

High quality sequence stop: 609.

FEATURES

Source

1.609

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="L11SNU35481-29-G02"

/sex="M"

/tissue type="Liver"

/cell_type="polygonal"

/cell_line="SNU-354"

/lab_host="Top10F"

/clone_id="L11SNU35481"

/note="Organ: Liver; Vector: pcNS-D2; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promotor as 5' primer and N(dT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

BASE COUNT 203 a 105 c 137 g 164 t

ORIGIN

Query Match 61.1%; Score 607.4; DB 14; Length 609;

Best Local Similarity 99.8%; Pred. No. 1.1e-114;

Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 45 CATATGAAAAATATGACAATAATGAAACACTTTGCAAAAATTGATGAAAAATCAT 104

Db 1 CATATGAAAAATATGACAATAATGAAACACTTTGCAAAAATTGATGAAAAATCAT 60

Qy 105 AAACATTCAGTTAATAACCGGTATTTCCAGTTCCAAGATGAATACAGTTAGATGAAGTG 164

Db 61 AAACATTCAGTTAATAACCGGTATTTCCAGTTCCAAGATGAATACAGTTAGATGAAGTG 120

Qy 165 ATGCATCTAAGAAGTTTGTATTTTGACTATCTTACAATGTTGTCCCACTTCAGAT 224

Db 121 ATGCATCTAAGAAGTTTGTATTTTGACTATCTTACAATGTTGTCCCACTTCAGAT 180

Qy 225 GGTGCTGCAGCAGCAATTTGGCCAGTGAAAGCATTTGTACAGAAATATGGCTGCATCC 284

Db 181 GGTGCTGCAGCAGCAATTTGGCCAGTGAAAGCATTTGTACAGAAATATGGCTGCATCC 240

Qy 285 AAAGCTGTGAAAATTTTGGCAAGAAGATGATGACTGATTTGCCAAGCTCGTTGAAGAA 344

Db 241 AAAGCTGTGAAAATTTTGGCAAGAAGATGATGACTGATTTGCCAAGCTCGTTGAAGAA 300

Qy 345 AAAGCATTTAATAATGTTGGCTTTGATATGAGTAAGAAGCTGCAAGAAAATGCTAT 404

Db 301 AAAGCATTTAATAATGTTGGCTTTGATATGAGTAAGAAGCTGCAAGAAAATGCTAT 360

Qy 405 GAGAAATCTGGCTGACACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCT 464

Db 361 GAGAAATCTGGCTGACACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCT 420

Qy 465 ACCAAGCAACTCCTTACTTATGAAGCACTGGGACTCTGTCCAGAAAGAGCAAGGTGCAAG 524

Db 421 ACCAAGCAACTCCTTACTTATGAAGCACTGGGACTCTGTCCAGAAAGAGCAAGGTGCAAG 480

```

QY 525 CTGTTGATAGAGAGATATACATATGAGGAAAGTGGTCATTAATCCTAGTGTGGA 584
    |||||||
Db 481 CTGTTGATAGAGAGATATACATATGAGGAAAGTGGTCATTAATCCTAGTGTGGA 540
QY 585 CTGATTTCAAAGGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAGAACTCTGC 644
    |||||||
Db 541 CTGATTTCAAAGGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAGAACTCTGC 600
QY 645 TGGCAGCTG 653
    |||||||
Db 601 TGGCAGCTG 609

RESULT 4
BE887512 600 bp mRNA linear EST 20-OCT-2000
LOCUS 601508161F1 NIH_MGC_71 Homo sapiens CDNA clone IMAGE:3909614 5',
DEFINITION mRNA sequence.
ACCESSION BE887512
VERSION BE887512.1 GI:10342878
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9723 row: j column: 15
High quality sequence stop: 598.

FEATURES
source location/Qualifiers
1..600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3909614"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 204 a 101 c 135 g 160 t
ORIGIN

Query Match 59.2%; Score 588; DB 10; Length 600;
Best Local Similarity 99.8%; Pred. No. 1.1e-110;
Matches 599; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 36 GGAAGAAGACATATGAAAAATATGCAACAAAAATTGAACACTTTGCCAAAAATTGATGG 95
    |||||||
Db 1 GGAAGAAGACATATGAAAAATATGCAACAAAAATTGAACACTTTGCCAAAAATTGATGG 60
QY 96 AAAAATCATAAACATTCAGTTAATAACCCGATATCCAGTTCAGATGAATACAGTTTA 155
    |||||||
Db 61 AAAAATCATAAACATTCAGTTAATAACCCGATATCCAGTTCAGATGAATACAGTTTA 120
QY 156 GATGAAGTATGGCATCTAAAGAGTTTGTATTTTGTACTATCTTACAATGTGTCCC 215
    |||||||
Db 121 GATGAAGTATGGCATCTAAAGAGTTTGTATTTTGTACTATCTTACAATGTGTCCC 180
QY 216 ACTTCAGATGGTGTGTCAGCAGCAATTTTGGCCAGTGAAGCAATTGTACAGAAGTATGGC 275
    |||||||

```

```

Db 181 ACTTCAGATGGTGTGTCAGCAGCAATTTTGGCCAGTGAAGCATTTGTACAGAAGTATGCC 240
QY 276 CTGCAATCCAAAGCTGTGAAATTTTGGCACACAGAATGATGACTGATTTGCCAAGCTCG 335
    |||||||
Db 241 CTGCAATCCAAAGCTGTGAAATTTTGGCACACAGAATGATGACTGATTTGCCAAGCTCG 300
QY 336 TTGGAAGAAAAAGCATTTAATAATGTTGGCTTTGATATGAGTAAGAGCTGCAGA 395
    |||||||
Db 301 TTGGAAGAAAAAGCATTTAATAATGTTGGCTTTGATATGAGTAAGAGCTGCAGA 360
QY 396 AAATGCTATGAGAAATCTGGCCTGACACCAAAATGATATTGACGTAATAGAACTTCAGAT 455
    |||||||
Db 361 AAATGCTATGAGAAATCTGGCCTGACACCAAAATGATATTGACGTAATAGAACTTCAGAT 420
QY 456 TGCTTTTCTACCAACGAACCTCTTACTTATGAAGCACTGGGACTCTGTCCAGAAGCAAA 515
    |||||||
Db 421 TGCTTTTCTACCAACGAACCTCTTACTTATGAAGCACTGGGACTCTGTCCAGAAGCAAA 480
QY 516 GGTGCAACGCTGTTGATAG-AGAGATATACATATGAGGAAAGTGGTCATAAATCC 574
    |||||||
Db 481 GGTGCAACGCTGTTGATAGAGAGATATACATATGAGGAAAGTGGTCATAAATCC 540
QY 575 TAGTGTGACTGATTTCAAAGGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGC 634
    |||||||
Db 541 TAGTGTGACTGATTTCAAAGGGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGC 600

RESULT 5
BC029332
LOCUS BC029332 1969 bp mRNA linear HTC 01-MAY-2002
DEFINITION Homo sapiens, clone IMAGE:5109488, mRNA.
ACCESSION BC029332
VERSION BC029332.1 GI:20379648
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC Project URL: http://mgc.nci.nih.gov
AUTHORS Contact: MGC help desk
TITLE Email: cgapbs-r@mail.nih.gov
JOURNAL Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAK Plate: 42 Row: m Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein
This clone has the following problem: no 5' EST match.

FEATURES
source location/Qualifiers
1..1969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5109488"

```

/clisue_type="Cervix", carcinoma"
/clone_lib="NIH_MGC_12"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

BASE COUNT 545 a 417 c 537 g 470 t

ORIGIN

Query Match		58.0%;	Score 576.6;	DB 11;	Length 1969;
Best Local Similarity		79.1%;	Pred. No. 2.4e-108;		
Matches 767;		Conservative 0;	Mismatches 184;	Indels 19;	Gaps 7;
QY	1	CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGGAAGAACATATGAAAAATATG	60		
Db	485	CACATCCGATTACTCCTCAGATGTTGGGTATGCTGGGAAGAACATATGAAAAATATG	544		
QY	61	GAACAAAAATTGAACACTTTGCAAAAAATTGGATGAAAAATCATAAACATTCACTTAATA	120		
Db	545	GAACAAAAAGTTGAACACTTTGCAAAAAATTGGATGAAAAATCATAAACACTCAGTTAATA	604		
QY	121	ACCCGTATTCCAGTCCCAAGATGAATACAGTTAGATGAAGTGATGCCATCTAAAGAG	180		
Db	605	ACAGTATTCCAGTCCCAAGATGAATACAGCTTAGAAGAAATTAATGAATCAAAACAG	664		
QY	181	TTTTGATTTTTTGACTATCTTACAATGTTGTCCCACTTCAGATGCTGCAGACGAA	240		
Db	665	TTTTCGATTTTCTGACTATCTTGAATGCTGCCCCACCTCAGATGCTGCCGTCGCCGA	724		
QY	241	TTTTGGCCAGTGAAGCATTGTGACAGAAGTATGCGCTGCAATCCAAAGCTGTGAAATTT	300		
Db	725	TTCTGTCCAGCGAGAGGTTGTGACAGCAGTACGCGCTGCAGTCCAAAGCGGTGAGATTG	784		
QY	301	TGGCACAAGAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAAAA	360		
Db	785	TGGCCCAAGAGATGATGACTGACTTACCAGTACGTTTGAAGAAAAAGTATTATTAAAG	844		
QY	361	TGGTTGGCTTGTATGAGTAAGAAGCTGCAGAAGAAATGCTATGAGAAATCTGCGCTGA	420		
Db	845	TGGTTGGCTATGATATGAGTAAGAAGCTGCAGAGAGATGCTATGAGAAGTCCGCGCTGA	904		
QY	421	CACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAACCTCTTA	480		
Db	905	CACCAACGATGTCAGCTGATAGAGCTTCACGATTGCTTCTGTCAATGAACCTATCA	964		
QY	481	CTTATGAAGCACTGGAGCTGTCTCCAGAAGACAAAGGTGCAACGCTGTTGATAGAGAG	540		
Db	965	CTTACGAAGCACTGGGCGCTGTCTCAGAAGACAAAGGTGAACCTGTGTGACAGAGGGG	1024		
QY	541	ATAATACATATGAGGAAAGTGGGTCTATAATCTAGTGTGCACTGATTTCAAGGGAC	600		
Db	1025	ACAACACTTACGGAGGAAAGTGGGTCTATCAACCTAGTGAGGCTCATCTCAAGGGAC	1084		
QY	601	ACCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAAGACTTGTCTGGCAGCTGAGAGGGG	660		
Db	1085	ACCACTGGGCGCAACAGGCTGTGCTCAGTGTGCGGAGCTCTGTGGCAGCTGAGAGGGG	1144		
QY	661	AAGCCGAAAAAGAGCGAAAGTTCCTGTGCAAAAGGTGGCTCTGCNGCATTAATTANGCAT	720		
Db	1145	AAGCCGAAAA--GAGCAGGTTCCCGGGCAAAAGTGGCTCTGCAACAATCTAGGCTT	1202		
QY	721	TGGAGAACTGTGTTGTAACACTTACAAGATGGGTTTTCCCGAAGCGCCAGTTCC	780		
Db	1203	CGGAGAGCTGTGTTGTCAACCTCTACAGATGGG--TTTTCCCGAAGCTGCCAGCTCC	1260		
QY	781	TTTTAGAACTCATCAAAATTGAACCCNGTTCCAACCAAGCTCTGCAAGTNATNGTTAA	840		
Db	1261	TTTCAGGA--CGACCAAGTTTCAGCTGCTCCCA--CCAGCTCTGCAAGGGGATGATTCAA	1316		
QY	841	NGNAAAAATCTNGTTTAAAGNGGATTGAGAAGGAAACNTNAAAAGGGAANGGGGACA	900		
Db	1317	GGCAAACTCTGCTTTAA--GGAGATTGAGAAGAAAGCTTGA---AGAGGAAGGGGGAACA	1370		
QY	901	ATTTGTGAAAAAAATNCGGNGGAATTTTGGCCCTTCAAGGGGAAANAATGGCCCTGG	960		

Db	1371	GTTCTGAAGAAAAATCGGTGGC-----ATTTTGTGCTTCAAAAGTGAAGATGGCCCTGGA	1425
QY	961	GGCGTAAAG 970	
Db	1426	GGCAAGAAG 1435	

RESULT 6	AK004860	1974 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS					
DEFINITION	Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:130002N20 product:sterol carrier protein 2, liver, full insert sequence.				
ACCESSION	AK004860				
VERSION	AK004860.2	GI:26089914			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arawaka, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bul, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-Oka, K., Wang, K. H., Weltz, C., Whitaker, C., Wilming, L., Wyszaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085660				
PUBMED	11217851				
REFERENCE	5				
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research				

TITLE Group Phase I & II Team.
JOURNAL Analysis of the mouse transcriptome based on functional annotation
REFERENCE of 60,770 full-length cDNAs
AUTHORS Nature 420, 563-573 (2002)
6 (bases 1 to 1974)

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishi,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT On Dec 6, 2002 this sequence version replaced gi:12836362.
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGCGGCCGCACTCGAGTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence[5'
GAGAGAGAGAGATCCAGAGCTCAATTAAATTAAACCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOLR.

FEATURES
source location/Qualifiers

1..1974
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:1300002N20"
/db_xref="MGI:1895024"
/db_xref="taxon:10090"
/clone="1300002N20"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1974
misc_feature
/note="sterol carrier protein 2, liver (MGD|MGI:98254)"
/db_xref="MGI:98254"

BASE COUNT 533 a 544 g 477 t
ORIGIN

Query Match 58.0%; Score 576.6; DB 11; Length 1974;
Best Local Similarity 79.1%; Pred. No. 2.4e-108;
Matches 767; Conservative 0; Mismatches 184; Indels 19; Gaps 7;
QY 1 CTCACCCAGTGTCTCTCAGATGTTTGGGTATGCTGGAAGAACATATGAAAAATATG 60
Db 494 CACATCCGATTACTCTCAGATGTTGGGTATGCTGGGAAGACATATGAAAAATATG 553
QY 61 GAACAAAATGAACACTTTGCAAAATTTGGATGAAAAATCATAAACATTCAATTATA 120
Db 554 GAACAAAAGTTGAACACTTTGCAAAATTTGGATGAAAAATCATAAACACTCAATTATA 613
QY 121 ACCCGTATTCCTCAGTTCAGAGATGATACAGTTAGATGAAGTATGATGATCAAGAG 180

Db 614 ACAAGTATTCACAGTTCAGAGATGATACAGCTTAGAAGAATATGAATCAAAACCAG 673
QY 181 TTTTGAATTTTGTGACTATCTTACAATGTTGTGCCACTTCAGATGTGCTGCAGACGAA 240
Db 674 TTTTCGATTTTCTGACTATCTTCCAATGCTGCCCCCACTTCAGATGTGCGCTGCCGCA 733
QY 241 TTTTGGCCAGTGAAGCATTTGTACAGAAGTATGGCCCTGCAATCCAAAGCTGTGAAATTT 300
Db 734 TTCTGTCACGAGAGAGTTGTGTGACGACAGTACGGCTGTGAGTCCAAAGCGGTGAGATTG 793
QY 301 TGGCAAGAAGATGATGACTGATTGCCCAGCTCGTTTGAAGAAAAAGCATATTATAAA 360
Db 794 TGGCCAGAGATGATGACTGACTTACCAGTAGCTTTGAAGAAAAAGTATTATTAAAG 853
QY 361 TGGTTGGCTTGTATGATGATAAGAAAGCTGCAAGAAAAATGCTATGAGAAATCGGCTGA 420
Db 854 TGGTTGGCTATGATATGATGATAAGAAAGCTGCCAGAGATGCTATGAGAAGTCCGGCTGA 913
QY 421 CACCAATGATATGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACTCCTTA 480
Db 914 CACCCAAGCATGTGACGCTGATAGAGCTTCACGATTGCTCTGTCAATGAACTCATCA 973
QY 481 CTATGAAAGCACTGGGACTCTGTCCAGAAGACAAAGGTGCAACGCTGTTGATAGAGAG 540
Db 974 CTTACGAAGCACTGGGGCTGTGTCCAGAAGACAAAGGTGGAACCTGTGTGACAGAGGG 1033
QY 541 ATAATACATATGAGGAAGTGGGTATTAATCTTACTGATGCTGATTTCAAAGGAGC 600
Db 1034 ACAACACTTACGAGGAAGTGGGTATCAACCTTAGAGAGGCTCATCTCAAGGAGC 1093
QY 601 ACCCACTAGGCGCTACAGCTTGTCTCAGTGTGAGAACTCTGCGCAGCTGAGAGGG 660
Db 1094 ACCCACTGGGCGCAACAGCTGTGCTCAGTGTGCGGAGCTCTGCGCAGCTGAGAGGG 1153
QY 661 AAGCCGAAAAGAGGCAAGTCTCTGTGTGCAAAAGTGGCTCGNGCATATTTANGCAT 720
Db 1154 AAGCCGAAA--GAGGCAAGTCTCCGGGCAAAAGTGGCTGTGACACAACTCTAGGCT 1211
QY 721 TGGAGAACTGTGTTGTAACACTCTACAGATGGGTTTCCCGGAAGCCGCACTTCC 780
Db 1212 CGGAGGAGCTGTGTTGTACACCTCTACAGATGGG--TTTCCCGAAGCTGCCAGCTCC 1269
QY 781 TTTTGAACATCATCAAAATGGAAGCCNGTTCACCAAGCTCTGCAAGTATNGCTTAA 840
Db 1270 TTCAGGA--CGCACAGGTTTCAAGCTCTCCA--CCAGCTCTGACGGGATGATTTCAA 1325
QY 841 NGNAAAATCTNGTTTAAAGNGGATTGAGAAGGAAACNTNAAAGAGGGAAGGGGAACA 900
Db 1326 GGCAAACCTCGTCTTTAA--GAGATTGAGAAGAAAGCTTGA---AGAGGAAGGGAACA 1379
QY 901 ATTTGTGAAGAAAAATNGCGNGGAATTTTGGCCCTTCAAGGGAANAATGCGCTGG 960
Db 1380 GTTCGTGAAGAAAATCGTGGC-----ATTTTGGCTTCAAAAGTGAAGATGCGCTGGA 1434
QY 961 GGGGTAAAG 970
Db 1435 GGCAAGAAG 1444

RESULT 7
Bg289834 926 bp mRNA linear EST 21-FEB-2001
LOCUS 602384738F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4513688 5',
DEFINITION mRNA sequence.
ACCESSION Bg289834
VERSION Bg289834.1 GI:13046023
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 926)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10400 row: 1 column: 09
High quality sequence stop: 652.

FEATURES
Source Location/Qualifiers

1..926
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4513688"
/issue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_93"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 281 a 176 c 234 g 235 t

ORIGIN

Query Match 57.6%; Score 572.6; DB 10; Length 926;
Best Local Similarity 95.5%; Pred. No. 1.6e-107;
Matches 686; Conservative 0; Mismatches 20; Indels 12; Gaps 9;

OY 1 CTCACCCAGTGTCTCTCAGATGTTGGGTATGCTGAAAGAACATATGAAAAATATG 60
|||||
DB 59 CTCACCCAGTGTCTCTCAGATG-TTGGGTATGCTGAAAGAACATATGAAAAATATG 117
OY 61 GAACAAAATTGACACTTGCAGAAAATTTGGATGGAATAATCATAAACATTCAGTTAATA 120
|||||
DB 118 GAACAAAATTGACACTTGCAGAAAATTTGGATGGAATAATCATAAACATTCAGTTAATA 177
OY 121 ACCCGTATTCAGTCCAGATGAATACAGTTAGATGAAGTGATGGCATCTAAGAG 180
|||||
DB 178 ACCCGTATTCAGTCCAGATGAATACAGTTAGATGAAGTGATGGCATCTAAGAG 237
OY 181 TTTTGGATTTTGGACTATCTTACAATGTTGTGCCACTTCAGATGCTGCAGAGCAA 240
|||||
DB 238 TTTTGGATTTTGGACTATCTTACAATGTTGTGCCACTTCAGATGCTGCAGAGCAA 297
OY 241 TTTGGCCAGTGAAGCATTTGTACAGAGATAGGCTGCATCCAAAGCTGTGAAATT 300
|||||
DB 298 -TTGGCCAGTGAAGCATTTGTACAGAGATAGGCTGCATCCAAAGCTGTGAAATT 355
OY 301 TGGCACAAGAATGATGATGATTTGGCCAAGCTCGTTGAAGAAAAGCATTATTAATA 360
|||||
DB 356 TGGCACAAGAATGATGATGATTTGGCCAAGCTCGTTGAAGAAAAGCATTATTAATA 415
OY 361 TGGTGGCTTGTATGAGTAAGAAGCTGCAGAAAATGCTATGAGAAATCTGGCCTGA 420
|||||
DB 416 TGGTGGCTTGTATGAGTAAGAAGCTGCAGAAAATGCTATGAGAAATCTGGCCTGA 475
OY 421 CACCAAAATGATATTGACGTAATAGAACTTCAGATGCTTTTCTACCAAGCAACTCTTA 480
|||||
DB 476 CACCAAAATGATATTGACGTAATAGAACTTCAGATGCTTTTCTACCAAGCAACTCTTA 535
OY 481 CTTATGAAGCACTGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 540
|||||
DB 536 CTTATGAAGCACTGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 595
OY 541 ATAATACATATGAGGAAAGTGGTCATAAATCC--TAGTGGTGAAGTGAATTTCAA--GG 597
|||||
DB 596 ATAATACATATGAGGAAAGTGGTCATAAATCCTTAGTGGTGAAGTGAATTTCAAAGG 655

OY 598 GACACCCACTAGG--CGCTACAGGTCTTGTCTCAGTGTGCAGAACT-CTGCTGGCAGCTGA 654
|||||
DB 656 GACACCCACTAGGCGCGCTAACAGAGTCTTGTCTCAGTGTGCAGAACTCCTGTTGGGAGCTGA 715

OY 655 GA-GGGGAAGCCGGAAAAGAGCCAAAGTTCCTGGTGCAAAAGGTGGCTCTGCNGCATAA 711
|||||
DB 716 GAGGGGAAGCCCGAAAAGAGGCAA--GTCTGTGTGCCAAAGTGTCTGCGGCAAAATTA 771

RESULT 8
BG105684 943 bp mRNA linear EST 30-JAN-2001
LOCUS BG105684
DEFINITION 602312282F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4421732 5',
mRNA sequence.

ACCESSION BG105684
VERSION BG105684.1 GI:12599530
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS 1 (bases 1 to 943)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10161 row: d column: 21
High quality sequence start: 5
High quality sequence stop: 586.

FEATURES
Source Location/Qualifiers

1..943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4421732"
/issue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_84"
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 282 a 225 c 249 g 187 t

Query Match 56.1%; Score 557.4; DB 10; Length 943;
Best Local Similarity 97.0%; Pred. No. 2.1e-104;
Matches 620; Conservative 0; Mismatches 13; Indels 6; Gaps 5;

OY 129 TCCAGTTCGAAGATGATACAGTTTAGATGAAGTGATGGCATCTAAGAAGTTTGTGAT 188
|||||
DB 8 TCCAGTTCGAAGATGATACAGTTTAGATGAAGTGATGGCATCTAAGAAGTTTGTGAT 67
OY 189 TTTTGAATCTTACATGTTGTCCCACTTCAGATGCTGCTGCAGAGCAATTTGGCC 248
|||||
DB 68 TTTTGAATCTTACATGTTGTCCCACTTCAGATGCTGCTGCAGAGCAATTTGGCC 126
OY 249 AGTGAAGCATTTGTACAGAAAGTATGGCTGCATCCAAAGCTGTGAAATTTGGCACA 308
|||||
DB 127 AGTGAAGCATTTGTACAGAAAGTATGGCTGCATCCAAAGCTGTGAAA-TTTGGCACA 185
OY 309 GAAATGATGATGATTTGCCAAGCTCGTTGAAGAAAAGCATTATTAATGTTGGC 368
|||||

Db 186 GAAATGATGACTGATTGTCACAGCTCGTTTGAGAAAAAGACATTATTAATGTTGCC 245

QY 369 TTTGATATGAGTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCCTGACACCAAAAT 428

Db 246 TTTGATATGAGTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCCTGACACCAAAAT 305

QY 429 GATATGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAACTCCTTACTTATGAA 488

Db 306 GATATGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAACTCCTTACTTATGAA 365

QY 489 GCACTGGGACTCTGTCCAGAGACAGGTGCAACGCTGGTTGATAGAGAGATAATACA 548

Db 366 GCACTGGGACTCTGTCCAGAGACAGGTGCAACGCTGGTTGATAGAGAGATAATACA 425

QY 549 TATGAGAGAAAGTGGTCATAAATCTAGTGGTGAAGTGAATTCAAAGGACACCCACTA 608

Db 426 TATGAGAGAAAGTGGTCATAAATCTAGTGGTGAAGTGAATTCAAAGGACACCCACTA 485

QY 609 GCGCGTACAGGTCTTGTCTCAGTGTGAGAACTCTGTGCGAGCTGAGAGGGAAGCCGGA 668

Db 486 GCGCGTACAGGTCTTGTCTCAGTGTGAGAACTCTGTGCGAGCTGAGAGGGAAGCCGGA 544

QY 669 AAAGAGCAAAGTCTCTGTGTCAGAGGTGGCTCTGCNGCATTAATTANGCATTGAGAGAA 728

Db 545 AA--GAGGCAAGTTCCTGTGTGC--AAGTGGCTCTGCAAGCATTAATTAGGCAATTGAGAGAG 601

QY 729 CTGTGGTTGTAACACTCTTACCAAGATGGGGTTTCCCGGA 767

Db 602 CTGTGTGTGTAACACTCTTACCAAGATGGGGTTTCCCGGAA 640

RESULT 9
AV653303
LOCUS AV653303 668 bp mRNA linear EST 15-JAN-2002
DEFINITION AV653303 GLC Homo sapiens cDNA clone GLCDB04 3', mRNA sequence.
ACCESSION AV653303
VERSION AV653303.1 GI:9874317
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 668)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PUBMED 11752456
COMMENT Contact: Zeguangu Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1.668
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCDB04"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 207 a 116 c 176 g 168 t 1 others

ORIGIN

Query Match 53.6%; Score 532.4; DB 9; Length 668;
Best Local Similarity 92.7%; Pred. No. 3e-99;
Matches 608; Conservative 0; Mismatches 38; Indels 10; Gaps 5;

QY 229 CTGCAGCAGCAATTTGGCCAGTGAAGCATTTGTACAGAAATGATGGCTGCAATCCAAAG 288

Db 1 CTGCAGCAGCAATTTGGCCAGTGAAGCATTTGTACAGAAATGATGGCTGCAATCCAAAG 60

QY 289 CTGTGGAATTTTGGCAGACAGAAATGATGACTGATTTGCCAAGCTCGTTGAAGAAAAA 348

Db 61 CTGTGGAATTTTGGCAGACAGAAATGATGACTGATTTGCCAAGCTCGTTGAAGAAAAA 120

QY 349 GCATTATTAATGTTGGCTTTGATATGAGTAAGAACTGCAGAAAAATGCTATGAGA 408

Db 121 GCATTATTAATGTTGGCTTTGATATGAGTAAGAACTGCAGAAAAATGCTATGAGA 180

QY 409 AATCTGGCTGACACCAAATGATATGACGTAATAGAACTTCAGATTGCTTTCTACCA 468

Db 181 AATCTGGCTGACACCAAATGATATGACGTAATAGAACTTCAGATTGCTTTCTACCA 240

QY 469 AGCAACTCCTTACTTATGAGCACTGGACTCTGTCCAGAAAGACAAAGTGCACGCTGG 528

Db 241 AGCAACTCCTTACTTATGAGCACTGGACTCTGTCCAGAAAGACAAAGTGCACGCTGG 300

QY 529 TTGATAGAGAGATAATACATATGAGAGAAAGTGGTCATAATCCTAGTGTGACTGA 588

Db 301 TTGATAGAGAGATAATACATATGAGAGAAAGTGGTCATAATCCTAGTGTGACTGA 360

QY 589 TTTCAGAGGACACCCCACTAGCCGCTACAGGTCTTCTCAGTGTGACAGACTCTGCTGCC 648

Db 361 TTTCAGAGGACACCCCACTAGCCGCTACAGGTCTTCTCAGTGTGACAGACTCTGCTGCC 420

QY 649 AGCTGAGAGGGGAAGCCGGAAGAGAGCAAGTTCCTGTGTGCAAAAGTGGCTCTGCNCA 708

Db 421 AGCTGAGAGGGGAAGCCGGAAGAGAGCAAGTTCCTGTGTGCAAAAGTGGCTCTGCAGCA 478

QY 709 TAAATTANGCATTGGAGAACTGTGTTGTAACACTCTCAAGATGGGGTTTCCCGGAA 768

Db 479 TAAATTANGCATTGGAGAGCTGTGTTGTAACACTCTCAAGATGG--TTTCCCGGAA 536

QY 769 GCCGCCAGTCTCTTTTGAACCTCATCAAAATGAAGCCNGTTCCAACCAAGCTCTGCAAG 828

Db 537 GCCGCCAGTCTCTTTTGAACCTCATCANATTTGAA---GCTGTTCCAACCAAGCTCTGCAAG 592

QY 829 TNAATNGTTTANGNAAAATCTNGTTTAAAGNGATTTGAGAGGAAGAAACNTNAAA 884

Db 593 TGATGATTTTAAGGCAAAATCTTGTTTTAA--GGAGATTGAGAGAAACTGAAGA 646

RESULT 10
CB951253
LOCUS CB951253 757 bp mRNA linear EST 29-APR-2003
DEFINITION AGENCOURT 13445466 NIH MGC 177 Mus musculus cDNA clone
IMAGE:30316161 5', mRNA sequence.
ACCESSION CB951253
VERSION CB951253.1 GI:30206290
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 757)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM107 row: b column: 10
High quality sequence stop: 551.

FEATURES

source

1..757
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30316161"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_177"
/note="Organ: liver; Vector: pDNR-LIB; Site 1: SfiI
(ggccatcggcc); Site 2: SfiI (ggcgctcgcc); cDNA made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAACGACGAGTGGCCATTACGGCCGG-3' and
5'-ATTCTAGAGCGCCGAGCGCCGACATG-dT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC library."

BASE COUNT 212 a 174 c 210 g 160 t 1 others

ORIGIN

Query Match 51.7%; Score 514; DB 14; Length 757;
Best Local Similarity 83.8%; Pred. No. 1.8e-95;
Matches 605; Conservative 0; Mismatches 112; Indels 5; Gaps 2;

OY 50 GGAATAATATGGAACAAAAATGAAACACTTGCAGAAAAATGGATGAAAAATCATTAACA 109
|||
DB 3 GGAATAATATGGAACAAAAAGTTGAACACTTGCAGAAAAATGGATGAAAAATCATTAACA 62
OY 110 TTCAGTTAATAACCCGATATCCAGTTCAGAGATGAATACAGTTAGATGAAGTATGAGC 169
|||
DB 63 CTCAGTTAATAACAGTATCCAGTTCAGAGATGAATACAGTTAGAGAGATGAATGA 122
OY 170 ATCTAAAGAGTTTGTGATTTTGAATATCTTACAATGTTGTCCACTTCAGATGTGC 229
|||
DB 123 ATCAAAACAGTTTTCGATTTTCTGACTATCTTGAATGTGCCCCACCTCAGATGTGC 182
OY 230 TGCAGACGAATTTTGGCCAGTGAAGCATTGTGTACAGAAATGAGCTGCAATCCAAAGC 289
|||
DB 183 CGCTGCCCGGATTTCTGTCAGCAGAGAGTTGTGCAGCAGTACGGCCTGCAGTCCAAAGC 242
OY 290 TGTGGAATTTTGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAAGAAAAAG 349
|||
DB 243 GTTGAAGATTGTGGCCAGAGATGATGACTGACTTACCCAGTACGTTTGAAGAAAAAG 302
OY 350 CATTAATAAATGTTGGCTTGTATATGATTAAGAAAGCTGCAAGAAATGCTATGAGAA 409
|||
DB 303 TATTATTAAGTGTGGCTATGATATGATTAAGAAAGCTGCAAGAGATGCTATGAGAA 362
OY 410 ATCTGAGCTGACACCAATGATTAATGAGTAATAGAACTTCAAGATTGCTTTCTACCA 469
|||
DB 363 GTCCGGCTGACACCAAGATGTCAGCTGATAGAGCTTCAAGATTGCTTCTGTCA 422
OY 470 CGAAGCTCTTACTTATGAAGAGCTGGAGCTCTGTCCAGAGAGCAAGGTGCAAGCTGTG 529
|||
DB 423 TGAAGCTCATCTTACGAAGCAGTGGGCTGTGTCCAGAGAGCAAGGTGCAAGCTGTG 482
OY 530 TGATAGAGAGATATATACATATGAGAGAAAGTGGTCAATAATCCTAGTGTGACTGAT 589
|||
DB 483 GGACAGAGAGGAGCAACACTTACGAGAGAAAGTGGTCAATCAACCTAGTGAAGGCTCAT 542
OY 590 TTCAAAAGGACACCCCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAAGAACTCTGCTGCA 649
|||
DB 543 CTCAAAAGGACACCCCACTGGCGGCAACAGGTCTGGCTCAGTGGCGGAGAGCTCTGCTGCA 602
OY 650 GCTGAGAGGGGAGCCCGAAAAAGGCAAGGTTCTGTGTCAGAAAGGTGGCTCTGCAGCAT 709
|||

DB 603 GCTGAGAGCGAAGCCGGA--GAGCAGTTCCCGGGCAAAAGTGTGCTTCAGACAC 660
OY 710 AATTANGCATTTGAGGAAGTGTGTTGTAACACTTACA--AGATGGGTTTCCCGG 766
|||
DB 661 AATCTAGGCTCGAGAGAGCTGTGTTGTTCACCCCTTACANGAGATGGGTTTCCCGG 720

OY 767 AA 768
||
DB 721 AA 722

RESULT 11

CB950754

LOCUS

DEFINITION

AGENCOURT_13777269 NIH_MGC_177 Mus musculus cDNA clone

IMAGE:30322082 5', mRNA sequence.

CB950754

VERSION

CB950754.1 GI:30205281

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 866)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Michael Brownstein Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDCM122 row: 1 column: 03

High quality sequence stop: 481.

location/Qualifiers

1..866

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:30322082"

/lab_host="DH10B (T1-phage-resistant)"

/clone_lib="NIH_MGC_177"

/note="Organ: liver; Vector: pDNR-LIB; Site 1: SfiI

(ggccatcggcc); Site 2: SfiI (ggcgctcgcc); cDNA made

by oligo-dT priming and directionally cloned. 5' and 3'

adaptors were used in cloning as follows:

5'-AAGCAGTGGTATCAACGACGAGTGGCCATTACGGCCGG-3' and

5'-ATTCTAGAGCGCCGAGCGCCGACATG-dT(30)NN-3'. Full-length

enriched library was constructed using the Clontech

Creator SMART kit and size-selected to contain the 0.5 kb

size fraction. Library created in the laboratory of M.

Brownstein (NIMH, NIH). Note: this is a NIH_MGC library."

BASE COUNT 240 a 200 c 242 g 183 t 1 others

ORIGIN

Query Match 51.5%; Score 512; DB 14; Length 866;

Best Local Similarity 82.8%; Pred. No. 4.6e-95;

Matches 607; Conservative 0; Mismatches 123; Indels 3; Gaps 2;

OY 81 GCAAAATTTGATGAAAAATCATTAACATTCAGTTAATAACCCGATATCCAGTTCCAA 140
|||
DB 3 GCAAAATTTGATGAAAAATCATTAACACTCAGTTAATAACACGATATCCAGTTCCAA 62
OY 141 GATGAATACAGTTTATGATGAAGTATGCGATCTTAAGAACTTTTGTGATTTTGTGACTATC 200
|||
DB 63 GATGAATACAGTTTATGAGAAAGTATGAATCAAAACAGTTTTCGATTTTCTGACTATC 122
OY 201 TTCAATGTTGCTCCACTTCAGATGTGTGTCAGAGCAATTTTGGCCAGTGAAGCATTT 260
|||

Db 123 TTGCAATGCTGCCCCACCTCAGATGTTGCCGCTGCCGATTTCTGTCCAGCGAGAGTTT 182

QY 261 GTACAGAGTATGGCCCTGCAATCCAAAGCTGTGGAATTTTGGCAACAAGATGATGACT 320

Db 183 GTGCAGCAGTACGGCCCTGCAGTCCAAAGCGGTGAGATTGTGCCCAGAGATGATGACT 242

QY 321 GATTGCCAAGCTCGTTTGAAGAAAAAGCATTTATTAATGCTTGGCTTTGATATGAGT 380

Db 243 GACTTACCCAGTACGTTTGAAGAGAAAAATATTATTAAGTGGTTGGCTATGATATGAGT 302

QY 381 AAAGAAGCTGCAAGAAAATGCTATGAGAAATCTGGCCCTGACACCAATGATATTGACGTA 440

Db 303 AAAGAAGCTGCCAGAGATGCTATGAGAAAGTCCGCCCTGACACCAAGATGTCGACGTG 362

QY 441 ATAGAAGCTCAGATGCTTTTCTTCAACGAACTCCTTACTTATGAAGCACTGGGACTC 500

Db 363 ATAGAGCTTCAAGATTGCTTCTGTCAATGAAGTCACTCATCGAAGCACTGGGCTC 422

QY 501 TGTCCAGAGGACAAAGGTGCAACGCTGTTGATAGAGAGATAATACATATGAGAGAAAG 560

Db 423 TGTCCAGAGGACAAAGGTGCAACGCTGTTGATAGAGAGATAATACATATGAGAGAAAG 482

QY 561 TGGGTCAATAATCTAGTGTGAGTGAATTTCAAAAGGACACCACTAGGCGCTACAGT 620

Db 483 TGGGTATCAACCTTAGTGAAGGCTCTATCTCAAAAGGACACCACTAGGCGCAACAGT 542

QY 621 CTTGCTCAGTGTGCAAGAACTCTGTGCGAGCTGAGAGGGGAAAGCCGGAAGAGGCAAG 680

Db 543 CTGGCTCAGTGTGCGAGAGCTCTGTGCGAGCTGAGAGCGGAAGCCGGAAG - GAGGCAAG 600

QY 681 TTCTGTGTCAAAGGTGCTGTGCGAGCTGAGAGGGGAAAGCCGGAAGAGGCAAGTGTGTA 740

Db 601 TTCCCGGGCAAAAGTGGCTCTGACAGCAATCTAGGCTCGGAGGAGCTGTGTTGTC 660

QY 741 CACTCTACAAGATGGGGTTTTCGCGGAAGCCGCAAGTCTCTTTAGAACTCATCAAAAT 800

Db 661 ACCTCTACAGGATGGGTTTTC - GAAGCTGCCAGCTCCCTTCAGAGACGACCAAGTTT 719

QY 801 GAAGCCNGTTCCA 813

Db 720 TCAGCTGGCTCCA 732

RESULT 12

AI525668 914 bp mRNA linear EST 09-AUG-1999

LOCUS PT1.3_04 D09.r tumor1 Homo sapiens cDNA 5', mRNA sequence.

DEFINITION AI525668

ACCESSION AI525668

VERSION AI525668.1 GI:4439803

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Huang,G.M., Ng,W.L., Farkas,J., He,L., Liang,H.A., Gordon,D., Yu,J. and Hood,L.

TITLE Prostate cancer expression profiling by cDNA sequencing analysis

JOURNAL Genomics 59 (2), 178-186 (1999)

MEDLINE 99339982

PUBMED 10409429

COMMENT Contact: Guyang Matthew Huang

FEATURES

source 1..914

location/Qualifiers

1..914

/organism="Homo sapiens"

/mol_type="mRNA"

Department of Molecular Biotechnology, Box 357730, University of Washington, Seattle, WA 98195

Tel: 5106280100

Fax: 5106280108

Email: huanggm@yahoo.com.

/db_xref="taxon:9606"

/clone_id="tumor1"

/note="Organ: Prostate; Vector: pBluescript; Directional cDNA library was constructed using Lambda ZP II kit (Stratagene). mRNA was extracted from a frozen prostate tumor tissue (Mayo Clinics)."

BASE COUNT 258 a 174 c 255 g 189 t 38 others

ORIGIN

Query Match 51.2%; Score 508.8; DB 9; Length 914;

Best Local Similarity 86.3%; Pred. No. 2.1e-94;

Matches 581; Conservative 0; Mismatches 84; Indels 8; Gaps 3;

QY 326 GCCAAGCTCGTTTGAAGAAAAAGCATTTATTAATGTTGGCTTTGATATGAGTAAGA 385

Db 19 GCCAAGCTCGTTTGAAGAAAAAGCATTTATTAATGTTGGCTTTGATATGAGTAAGA 78

QY 386 AGCTGCAAGAAATGCTATGAGAAATCTGGCCCTGACACCAATGATATTGACGTAATGA 445

Db 79 AGCTGCAAGAAATGCTATGAGAAATCTGGCCCTGACACCAATGATATTGACGTAATGA 138

QY 446 ACTTCAAGATTGCTTTTCTACCAACGAACTCCTTACTTATGAAGCACTGGGACTGTCC 505

Db 139 ACTTCAAGATTGCTTTTCTACCAACGAACTCCTTACTTATGAAGCACTGGGACTGTCC 198

QY 506 AGAAGACAAGGTGCAACGCTGTTGATAGAGAGATAATACATATGAGAGAAAGTGGGT 565

Db 199 AGAAGACAAGGTGCAACGCTGTTGATAGAGAGATAATACATATGAGAGAAAGTGGGT 258

QY 566 CATTAATCTAGTGTGCACTGATTTCAAGGAGACACCACTAGGCGCTACAGTCTTGC 625

Db 259 CATTAATCTAGTGTGCACTGATTTCAAGGAGACACCACTAGGCGCTACAGTCTTGC 318

QY 626 TCAGTGTGCAAGCTCTGCTGCGACAGCTGAGAGGGGAAAGCCGGAAGAGGCAAGTTCCT 685

Db 319 TCAGTGTGCAAGCTCTGCTGCGACAGCTGAGAGGGGAAAGCCGGAAG - GAGGCAAGTTCCT 376

QY 686 GGTGCAAGGTGCTCTGCGCATTAATTANGCATTTGAGAGAACTGTGTTGTAACATC 745

Db 377 GGTGCAAGGTGCTCTGCGCATTAATTANGCATTTGAGAGAGCTGTGTTGTAACATC 436

QY 746 TACAAGATGGGTTTTCGCGGAAGCCGCAAGTCTTTTGAAGTCAATCAAAATGGAAGC 805

Db 437 TACAAGATGGG - TTTTCGCGGAGCCGCAAGTCTTTTGAAGTCTCTCAAAATGGAAGC 494

QY 806 CNGTTCACCAAGCTCTGCAAGTNATNGTTAANG --- NAAATCTNGTTTAAAGG 861

Db 495 TGTTCACCAAGCTCTGCAAGTATGATGATTTAAAGGCAAAACCTGTTTAAAGG 554

QY 862 NGGATTGAGAGGAACNTNAAAGAGGGAAGGGGAACAATTTGTGAAGAAAAATNCGG 921

Db 555 GGGTTGNMAAGAACTTTGAAAGAGGAANGGGGAACCAAGTTGTGNAGAAAAACCGG 614

QY 922 NGGGAATTTTTCCTTCAAGGGGAANAATGGCCCTGGGGGTTAAAGANGGCCACCT 981

Db 615 GGNATTTTNNCCCTCCAAAGNGGAAGATGGCCCTGGGGGGGAAAAAGGCCNCCCT 674

QY 982 GGGGTGTGGAT 994

Db 675 GGGGTGTGGAT 687

RESULT 13

AL694422 583 bp mRNA linear EST 21-MAR-2002

LOCUS AL694422

DEFINITION DKFZp313F1641_r1 313 (synonym: h1cc2) Homo sapiens cDNA clone

ACCESSION AL694422

VERSION AL694422.1 GI:19617879

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 583)	Koehrer, K., Beyer, A., Mewes, W., Well, B. and Wiemann, S.	EST (Koehrer, K., Beyer, A., Mewes, H.W., Well, B. and Wiemann, S.)	Unpublished	Contact: Koehrer K

BASE COUNT ORIGIN	199 a	97 c	121 g	166 t
----------------------	-------	------	-------	-------

RESULT 14	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AL694509	577 bp mRNA	linear	EST 21-MAR-2002									
AL694509	DKFZp313G1541_r1_313 (synonym: hlcc2)	Homo sapiens	CDNA clone	DKFZp313G1541_5', mRNA sequence.								
AL694509				AL694509								
AL694509.1	GI:19617887			EST.								
				Homo sapiens (human)								
				Homo sapiens								
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
				1 (bases 1 to 577)								
				Koehrer, K., Beyer, A., Mewes, W., weil, B. and Wiemann, S.								
				EST (Koehrer, K., Beyer, A., Mewes, H.W., weil, B. and Wiemann, S.)								
				Unpublished								
				Contact: Koehrer K								

BASE COUNT	ORIGIN
196 a	97 c
119 g	165 t
	.

Db 441 ACACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACCTCCTT 500
Qy 480 ACTTATGAGCACTGGAGACTCTGTCCAGAAGACAAAGTGCAACGCTGTTGATAGAGA 539
Db 501 ACTTATGAGCACTGGAGACTCTGTCCAGAAGACAAAGTGCAATGCTGTTGATAGAGA 560
Qy 540 GATAATACATATGGAG 556
Db 561 GATAATACATATGGAG 577

RESULT 15

BM552173 1051 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6543441 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5549545
DEFINITION 5', mRNA Sequence.

ACCESSION BM552173
VERSION BM552173.1 GI:18789814
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12260 row: e column: 02
High quality sequence stop: 653.

FEATURES

source

1.1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5549545"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

BASE COUNT 318 a 189 c 258 g 280 t 6 others
ORIGIN

Query Match 47.3%; Score 470.2; DB 12; Length 1051;
Best Local Similarity 96.4%; Pred. No. 1.8e-86;
Matches 502; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

Qy 1 CTCACCCAGTTGCTCCTCAGATGTTGGGTATGCTGGAAGAAACATATGGAATAATATG 60
Db 307 CTCACCCAGTTGCTCCTCAGATGTTGGGTATGCTGGAAGAAACATATGGAATAATATG 366
Qy 61 GAACAAAATTTGAACACTTTGCAAAAATTTGATGAAAAATCATAAACATTCAAGTTAATA 120
Db 367 GAACAAAATTTGAACACTTTGCAAAAATTTGATGAAAAATCATAAACATTCAAGTTAATA 426
Qy 121 ACCCGTATCCAGTCCAGAGATGATACAGTTAGATGAAGTATGCGATCTAAAGAAG 180
Db 427 ACCCGTATCCAGTCCAGAGATGATACAGTTAGATGAAGTATGCGATCTAAAGAAG 486
Qy 181 TTTTGAATTTTGAATCTTACATGTTGCTCCACTTCAGATGCTGCGACAGCAA 240
Db 487 TTTTGAATTTTGAATCTTACATGTTGCTCCACTTCAGATGCTGCGACAGCAA 546

Qy 241 TTTTGGCCAGTGAAGCAATTTGTACAGAAGTATGGCTGCAATCCAAAGCTGTGAAATTT 300
Db 547 TTTTGGCCAGTGAAGCAATTTGTACAGAAGTATGGCTGCAATCCAAAGCTGTGAAATTT 606
Qy 301 TGGCAACAAGAAATGATGATGATTTGGCCAGCTCGTTGAAGAAAAAGCATTTATTAATA 360
Db 607 TGGCAACAAGAAATGATGATGATTTGGCCAGCTCGTTGAAGAAAAAGCATTTATTAATA 666
Qy 361 TGGTTGGCTTTGATATGAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCTGA 420
Db 667 TGGTTGGCTTTGATATGAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCTGA 726
Qy 421 CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACCT-CTT 479
Db 727 CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACCTCCTT 786
Qy 480 ACTTATGAAGCACTGGACTCTGT---CCAGAAGGACAAG 517
Db 787 ACTTATGAAGCACTGGAGACTCTGTCTGCCAGAGACAAG 827

Search completed: November 27, 2003, 12:28:18
Job time : 2685.01 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:01:24 ; Search time 326.626 Seconds
(without alignments)
8215.019 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 994
Sequence: 1 ctcaccacgtgtgtctctccag.....ccaccctggggtgtggat 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	978	98.4	994	20	AAX40096 Gastric cancer ass
2	769	77.4	2663	22	AAH57501 Human liver cell s
3	769	77.4	3053	23	ABV21847 Human prostate exp
4	769	77.4	3053	23	ABV27671 Human prostate exp
5	573.4	57.7	2152	24	AB199762 Mouse ischaemic co
6	322.8	32.5	432	25	ABX40513 Bovine EST associa
7	266	26.8	1837	23	ABL16479 Drosophila melanog
8	260	26.2	1237	23	ABL16661 Drosophila melanog

9	253	25.5	346	25	ABX42390 Bovine EST associa
10	238	23.9	314	25	ABX35482 Bovine EST associa
11	214	21.5	4411529	22	AAI99682 Mycobacterium tube
12	211.4	21.3	3613	23	ABL16660 Drosophila melanog
13	211.4	21.3	4082	23	ABL18868 Drosophila melanog
14	211.4	21.3	4082	23	ABL26826 Drosophila melanog
15	208.2	20.9	4356	23	ABL16478 Drosophila melanog
16	199.6	20.1	4403765	22	AAI99683 Mycobacterium tube
17	189.8	19.1	412	25	ABX37781 Bovine EST associa
18	106	10.7	330	21	AAC07977 Human secreted pro
19	99.2	10.0	199	21	AAC07978 Human secreted pro
20	93.2	9.4	192	15	AAQ77270 Human genome fragm
21	92	9.3	249	25	ABX26831 Human GDP-mannose
22	90.4	9.1	1664976	19	AAV21209 Methanococcus jann
23	71.8	7.2	382	23	ABV12035 Human prostate exp
24	71.8	7.2	416	23	ABV33180 Human prostate exp
25	70.6	7.1	349980	22	AAH41226 Pyrococcus abyssi
26	53.2	5.4	705	24	ABT09578 phase-1 Rat CT gen
27	52.6	5.3	451	23	ABV42104 Human prostate exp
28	41.6	4.2	1230	10	AAN90223 Malaria-specific D
29	39.8	4.0	1662	18	AAT75075 DNA encoding argin
30	39.8	4.0	6313	18	AAV74512 Staphylococcus aur
31	39	3.9	580073	18	AAT58840 Mycoplasma genital
32	38.8	3.9	19387	22	AAS34561 Human DNA for a no
33	38.6	3.9	9207	22	AAS46789 Tumour suppressor
34	38.6	3.9	9207	24	ABK28433 DNA transcripction
35	38.4	3.9	327	23	ABV06120 Human prostate exp
36	38.4	3.9	2557	23	ABL15768 Drosophila melanog
37	38.4	3.9	3965	23	ABL02948 Drosophila melanog
38	38.4	3.9	5189	23	ABL25648 Drosophila melanog
39	38.2	3.8	1659	23	AAS51860 Staphylococcus aur
40	38.2	3.8	1662	23	AAS54694 Staphylococcus aur
41	38	3.8	10279	24	ABL92276 Chemically treated
42	38	3.8	10279	24	ABL33590 Human immune syste
43	38	3.8	10279	24	AAD22327 Chemically treated
44	37.8	3.8	1353	21	AAAS1780 S. cerevisiae esse
45	37.8	3.8	1664976	19	AAV21209 Methanococcus jann

ALIGNMENTS

RESULT 1	
ID	AAAX40096
ID	AAAX40096 standard; DNA; 994 BP.
XX	
AC	AAAX40096;
XX	
DT	02-JUL-1999 (first entry)
XX	
DE	Gastric cancer associated gene.
XX	
KW	Cancer associated antigen; diagnosis; research; treatment; human;
KW	breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW	prostate cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO9904265-A2.
XX	
PD	28-JAN-1999.
XX	
PF	15-JUL-1998; 98WO-US14679.
XX	
PR	22-JUN-1998; 98US-0102322.
PR	17-JUL-1997; 97US-0896164.
PR	10-OCT-1997; 97US-0061599.
PR	10-OCT-1997; 97US-0061765.
PR	10-OCT-1997; 97US-0948705.
PR	11-OCT-1997; 97GB-0021697.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.
XX	

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfeundschn M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
PS Claim 67; Page 695-696; 787pp; English.
XX
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
SQ Sequence 994 BP; 313 A; 168 C; 249 G; 248 T; 16 other;

Query Match 98.4%; Score 978; DB 20; Length 994;
Best Local Similarity 100.0%; Pred. No. 5e-266;
Matches 994; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGGAAAAAGACATATGAAAAATATG 60
Db 1 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGGAAAAAGACATATGAAAAATATG 60
QY 61 GAACAAAAATTGAACACTTTGCAAAAATTGGATGAAAAATCATAAACATTCACTTAATA 120
Db 61 GAACAAAAATTGAACACTTTGCAAAAATTGGATGAAAAATCATAAACATTCACTTAATA 120
QY 121 ACCCGTATCCAGTTCGAAGATGATACAGTTAGATGAAGTATGCGATCTAAAGAG 180
Db 121 ACCCGTATCCAGTTCGAAGATGATACAGTTAGATGAAGTATGCGATCTAAAGAG 180
QY 181 TTTTGGATTTTGGATCTTACATGTTGTCCACTTCAGATGGTGTGACAGACAA 240
Db 181 TTTTGGATTTTGGATCTTACATGTTGTCCACTTCAGATGGTGTGACAGACAA 240
QY 241 TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCTGCAATCCAAAGCTGTGAAATT 300
Db 241 TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCTGCAATCCAAAGCTGTGAAATT 300
QY 301 TGGCACAAGAAATGATGACTGATTGCCAAGCTCGTTGAAGAAAAAGCATTATTAAAA 360
Db 301 TGGCACAAGAAATGATGACTGATTGCCAAGCTCGTTGAAGAAAAAGCATTATTAAAA 360
QY 361 TGGTGGCTTGTATGATGATAAGAGCTGCAAGAAAAATGCTATGAGAAATCTGGCCTGA 420
Db 361 TGGTGGCTTGTATGATGATAAGAGCTGCAAGAAAAATGCTATGAGAAATCTGGCCTGA 420
QY 421 CACCAATGATATTGACGTAATGAACCTTACGATGCTTTTCTACCAACGAACCTCTTA 480
Db 421 CACCAATGATATTGACGTAATGAACCTTACGATGCTTTTCTACCAACGAACCTCTTA 480
QY 481 CTTATGAAGCACTGGACTCTGTCCAGAAGACAAAGTGCACGCTGCTGATAGAGAG 540
Db 481 CTTATGAAGCACTGGACTCTGTCCAGAAGACAAAGTGCACGCTGCTGATAGAGAG 540
QY 541 ATATATCATATGAGGAAAGTGGTGCATTAATCTAGTGTGAGCTGATTTCAAAGGAC 600
Db 541 ATATATCATATGAGGAAAGTGGTGCATTAATCTAGTGTGAGCTGATTTCAAAGGAC 600

QY 601 ACCCACTAGGCGCTACAGGTTCTGCTCAGTGTGACAACTCTGCTGCAGCTGAGAGGG 660
Db 601 ACCCACTAGGCGCTACAGGTTCTGCTCAGTGTGACAACTCTGCTGCAGCTGAGAGGG 660
QY 661 AAGCCGAAAGAGAGCAAGTTCCTGCTGCAAGGTGGCTCTGCNGCATATTTANGCAT 720
Db 661 AAGCCGAAAGAGAGCAAGTTCCTGCTGCAAGGTGGCTCTGCNGCATATTTANGCAT 720
QY 721 TGGAGGAACGTGTGTTGTAACACTTACAGATGGGGTTTCCCGGAAGCCGCAATTCC 780
Db 721 TGGAGGAACGTGTGTTGTAACACTTACAGATGGGGTTTCCCGGAAGCCGCAATTCC 780
QY 781 TTTTGAACCTCATCAAAATTGAAGCCNGTTCCAACCAAGCTCTGCAAGTATNGTTTAA 840
Db 781 TTTTGAACCTCATCAAAATTGAAGCCNGTTCCAACCAAGCTCTGCAAGTATNGTTTAA 840
QY 841 NGNAAAAATCTNGTTTAAAGNGGATTGAGAAGGAAACNTNAAAGAGGGAANGGGAACA 900
Db 841 NGNAAAAATCTNGTTTAAAGNGGATTGAGAAGGAAACNTNAAAGAGGGAANGGGAACA 900
QY 901 ATTGTGAAGAAAAATNCGNGGGAATTTTGGCCCTCAAGGGAANAATGGCCCTGG 960
Db 901 ATTGTGAAGAAAAATNCGNGGGAATTTTGGCCCTCAAGGGAANAATGGCCCTGG 960
QY 961 GGGGTAAAGANGGCCACCTGGGGTGTGGAT 994
Db 961 GGGGTAAAGANGGCCACCTGGGGTGTGGAT 994

RESULT 2
AAH57501
ID AAH57501 standard; cDNA; 2663 BP.

XX AC AAH57501;
XX DT 10-SEP-2001 (first entry)
XX DE Human liver cell specific cDNA sequence SEQ ID NO:341.
XX KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
OS Homo sapiens.
XX PN WO200132927-A2.
XX PD 10-MAY-2001.
XX PF 02-NOV-2000; 2000WO-US30396.
XX PR 04-NOV-1999; 99US-0163508.
XX PA (INCY-) INCYTE GENOMICS INC.
PI Sornasse T, Seilhamer JU, Watson GA;
DR WPI; 2001-291057/30.
XX
XX PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
XX
PS Claim 1; Page 258-259; 327pp; English.
XX
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and

CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.

XX Sequence 2663 BP; 807 A; 472 C; 599 G; 785 T; 0 other;

Query Match 77.4%; Score 769; DB 22; Length 2663;

Best Local Similarity 91.3%; Pred. No. 9.1e-207;

Matches 899; Conservative 0; Mismatches 66; Indels 20; Gaps 8;

QY 1 CTCACCCAGTGTCTCCCTCAGATGTTGGGTATGCTGAAAAAGAACATATGAAAAATATG 60
DB 566 CTCACCCAGTGTCTCCCTCAGATGTTGGGTATGCTGAAAAAGAACATATGAAAAATATG 625
QY 61 GAACAAAAATTGAACACTTTGCAAAAATTGGATGAAAAATCATAAACATTCACTTAATA 120
DB 626 GAACAAAAATTGAACACTTTGCAAAAATTGGATGAAAAATCATAAACATTCACTTAATA 685
QY 121 ACCCGTATTTCCAGTTCCTCAAGATGAATACAGTTTAGATGAAGTATGGCATCTAAGAG 180
DB 686 ACCCGTATTTCCAGTTCCTCAAGATGAATACAGTTTAGATGAAGTATGGCATCTAAGAG 745
QY 181 TTTTGTATTTTGTACTATCTTACAAATGTTGTCCCACTTCAGATGGTGTGACAGCA 240
DB 746 TTTTGTATTTTGTACTATCTTACAAATGTTGTCCCACTTCAGATGGTGTGACAGCA 805
QY 241 TTTTGGCCAGTGAAGCATTGTGACAGAAATGATGCTGCAATCCAAAGCTGTGAAATTT 300
DB 806 TTTTGGCCAGTGAAGCATTGTGACAGAAATGATGCTGCAATCCAAAGCTGTGAAATTT 865
QY 301 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAA 360
DB 866 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAA 925
QY 361 TGGTGGCTTTGATATGATTAAGAAAGCTGCAGAAAAATGCTATGAGAAATCTGGCTGA 420
DB 926 TGGTGGCTTTGATATGATTAAGAAAGCTGCAGAAAAATGCTATGAGAAATCTGGCTGA 985
QY 421 CACCAATGATATGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACCTCTTA 480
DB 986 CACCAATGATATGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACCTCTTA 1045
QY 481 CTTATGAAGCACTGGGACTCTGTCCAGAAAGACAAAGGTGCAACGCTGTTATAGAGAG 540
DB 1046 CTTATGAAGCACTGGGACTCTGTCCAGAAAGACAAAGGTGCAACGCTGTTATAGAGAG 1105
QY 541 ATAAATACATATGAGGAAAGTGGCTCATAAATCCTAGTGTGCACTGATTTCAAAGGAC 600
DB 1106 ATAAATACATATGAGGAAAGTGGCTCATAAATCCTAGTGTGCACTGATTTCAAAGGAC 1165
QY 601 ACCCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAAGAACTGTGCGACGTGAGAGGG 660
DB 1166 ACCCACTAGGCGCTACAGGTCTGTCTCAGTGTGCAAGAACTGTGCGACGTGAGAGGG 1225
QY 661 AAGCCGGAAGAGGCAAGTCTCTGTGCAAAAGGTGGCTGTGCGCATTAATTANGCAT 720
DB 1226 AAGCCGGAAGAGGCAAGTCTCTGTGCAAAAGGTGGCTGTGCGCATTAATTANGCAT 1283
QY 721 TGGAGGAAGTGTGTTGTAACACTCTCAAGATGGGGTTTCCGGAAAGCCGACAGTTC 780
DB 1284 TGGAGGAAGTGTGTTGTAACACTCTCAAGATGGG--TTTCCGGAAGCCGACAGTTC 1340
QY 781 TTTTAGAAGTCTCAAAATTTGAAGCNGTTCCAAACCAAGCTGTGCAAGTATNGGTTAA 840
DB 1341 TTTTAGAAGTCTCAAAATTTGAAGCNGTTCCAAACCAAGCTGTGCAAGTATNGGTTAA 1397
QY 841 NGNAAAATCTNGTTTAAAGGNGATTTGAGAAGAAACNTNAAAAGGGAANGGGAACA 900

DB 1398 GCGCAATCTGTGTTTAA--GGAGATTGAGAGAAACTTGA----AGAGGAAGGGGAACA 1451
QY 901 ATTTGTGAAGAAAAATNCGNGGGAATTTTGGCCTTCAAGGGGAANAATGGCCCTGG 960
DB 1452 GTTTGTGAAGAAAAATCGGTGT----ATTTTGGCTTCAAGGTGAAG-ATGGCCCTGG 1505
QY 961 GGGGTAAAGANGGCCACCTGGGG 985
DB 1506 GGGTAAAGAGGCCACCTGGGTGTG 1530

RESULT 3

ABV21847

ID ABV21847 standard; cDNA; 3053 BP.

XX

AC ABV21847;

XX

DT 13-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 21838.

XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200160860-A2.

XX

PD 23-AUG-2001.

XX

PF 20-FEB-2001; 2001WO-US05171.

XX

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

PI Schlegel R, Endege WO, Monahan JE;

XX

DR WPI; 2001-662795/76.

XX

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer -

XX

PS Claim 1; Page 3703; 11750pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX

SQ Sequence 3053 BP; 1010 A; 572 C; 887 G; 580 T; 4 other;

XX

Query Match

Best Local Similarity 77.4%; Score 769; DB 23; Length 3053;

Matches 899; Conservative 0; Mismatches 66; Indels 20; Gaps 8;

```
QY 1 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 60
Db 588 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 647
QY 61 GAACAAAAATGAAACACTTTGCAAAAAATGGATGAAAAATCATTAACATTCACTTAATA 120
Db 648 GAACAAAAATGAAACACTTTGCAAAAAATGGATGAAAAATCATTAACATTCACTTAATA 707
QY 121 ACCCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGAATGCATCTAAGAAG 180
Db 708 ACCCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGAATGCATCTAAGAAG 767
QY 181 TTTTGAATTTTGTACTATCTTACAATGTGTGCCACTTCAGATGCTGCAGACGCA 240
Db 768 TTTTGAATTTTGTACTATCTTACAATGTGTGCCACTTCAGATGCTGCAGACGCA 827
QY 241 TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCTGCAATCCAAGCTGTGAAATTT 300
Db 828 TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCTGCAATCCAAGCTGTGAAATTT 887
QY 301 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAAAA 360
Db 888 TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAAAA 947
QY 361 TGGTGGCTTTGATATGATGAAGAAAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA 420
Db 948 TGGTGGCTTTGATATGATGAAGAAAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA 1007
QY 421 CACCAATGATATTGACGTAATAGAACTTCAAGATGTCTTTTCTACCAACGAAGCTCCTTA 480
Db 1008 CACCAATGATATTGACGTAATAGAACTTCAAGATGTCTTTTCTACCAACGAAGCTCCTTA 1067
QY 481 CTTATGAAGCACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTGATAGAGAG 540
Db 1068 CTTATGAAGCACTGGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTGATAGAGAG 1127
QY 541 ATAATACATATGAGGAAAGTGGGTCTATAATCCTAGTGTGAGTGAATTTCAAGGAG 600
Db 1128 ATAATACATATGAGGAAAGTGGGTCTATAATCCTAGTGTGAGTGAATTTCAAGGAG 1187
QY 601 ACCCACTAGGCGCTACAGTCTGTGCTCAGTGTGCAAGAACTCTGTGGCAGCTGAGAGGG 660
Db 1188 ACCCACTAGGCGCTACAGTCTGTGCTCAGTGTGCAAGAACTCTGTGGCAGCTGAGAGGG 1247
QY 661 AAGCCGGAAGAAAGAGCAAGTCTCTGTGCAAAAGGTGCTGCGNCGATATTTANGCAT 720
Db 1248 AAGCCGGAAGAA--GAGGCAAGTCTCTGTGCAAAAGGTGCTGCGNCGATATTTANGCAT 1305
QY 721 TGGAGGAATGTGTGTGTACACCTCTACAAGATGGGGTTTCCCGGAAGCCGCAAGTTC 780
Db 1306 TGGAGGAATGTGTGTGTGTACACCTCTACAAGATGGG--TTTCCCGGAAGCCGCAAGTTC 1362
QY 781 TTTTAGAATCTATCAAAATTTGAAGCGNGTTCACCAAGCTCTGCAAGTATNGGTTTAA 840
Db 1363 TTTTAGAATCTATCAAAATTTGA--GCTGTTCACCAAGCTCTGCAAGTATNGGTTTAA 1419
QY 841 NGNAAAATCTNGTTTTAAAGNGATGAGAAGGAAACNTNAAAGAGGGAANGGGGAACA 900
Db 1420 GGCATAATCTTGTTTTAA--GGAGATGAGAAGAAACTTGA---AGAGGAAGGGGAACA 1473
QY 901 ATTTGTGAAGAAAGAAATNCGNGGGAATTTTGGCCTTCAAGGGGAANAATGGCCCTGG 960
Db 1474 GTTTGTGAAGAAATCGGTGT-----ATTTTGGCCTTCAAGGTGAAG-ATGGCCCTGG 1527
QY 961 GGGGTAAAAGANGGCCACCCCTGGGG 985
Db 1528 GGGTAAAAGAGGCCACCTGGGTGGTG 1552
```

RESULT 4
ABV27671
ID ABV27671 standard; cDNA; 3053 BP.

```
XX AC ABV27671;
XX XX
DT 16-SEP-2002 (first entry)
XX XX
DE Human prostate expression marker cDNA 27662.
XX XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX XX
PS Claim 1, Page 5667; 11750pp; English.
XX XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX XX
SQ Sequence 3053 BP; 1010 A; 572 C; 887 G; 580 T; 4 other;
XX XX
Query Match 77.4%; Score 769; DB 23; Length 3053;
Best Local Similarity 91.3%; Pred. No. 9.6e-207;
Matches 899; Conservative 0; Mismatches 66; Indels 20; Gaps 8;
QY 1 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 60
Db 588 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 647
QY 61 GAACAAAAATGAAACACTTTGCAAAAAATGGATGAAAAATCATTAACATTCACTTAATA 120
Db 648 GAACAAAAATGAAACACTTTGCAAAAAATGGATGAAAAATCATTAACATTCACTTAATA 707
QY 121 ACCCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGAATGCATCTAAGAAG 180
Db 708 ACCCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGAATGCATCTAAGAAG 767
QY 181 TTTTGAATTTTGTACTATCTTACAATGTGTGCCACTTCAGATGCTGCAGACGCA 240
Db 768 TTTTGAATTTTGTACTATCTTACAATGTGTGCCACTTCAGATGCTGCAGACGCA 827
```

Db 768 TTTTGGATTTTGGACTATCTTACAAATGTTGTCCCACTTCAGATGTGCTGCAGACGAA 827
QY 241 TTTTGGCCAGTGAAGCATTTGTCAGAGAAGTATGCGCTGCAATCCAAAGCTGTGAAATTT 300
Db 828 TTTTGGCCAGTGAAGCATTTGTCAGAGAAGTATGCGCTGCAATCCAAAGCTGTGAAATTT 887
QY 301 TGGCACAAGAATGATGATGATTTTGGCCAAGCTCGTTTGAAGAAAAAGCATTTATTA 360
Db 888 TGGCACAAGAATGATGATGATTTTGGCCAAGCTCGTTTGAAGAAAAAGCATTTATTA 947
QY 361 TGGTTGGCTTTGATATGATTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGA 420
Db 948 TGGTTGGCTTTGATATGATTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGA 1007
QY 421 CACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCTTA 480
Db 1008 CACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCTTA 1067
QY 481 CTTATGAAGCACTGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 540
Db 1068 CTTATGAAGCACTGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 1127
QY 541 ATAATACATATGAGAGAAAGTGGTTCATAATCCTAGTGTGGAAGTATTTCAAAGGAC 600
Db 1128 ATAATACATATGAGAGAAAGTGGTTCATAATCCTAGTGTGGAAGTATTTCAAAGGAC 1187
QY 601 ACCCACTAAGCGCTACAGGTCTTGTCTCAGTGTGCAAGAACTCTGCTGGCAGCTGAGAGGG 660
Db 1188 ACCCACTAAGCGCTACAGGTCTTGTCTCAGTGTGCAAGAACTCTGCTGGCAGCTGAGAGGG 1247
QY 661 AAGCCGGAAGAGGCAAGTCTCTGCTGCAAAAGGTGGCTCTGCNCGCATTAATTNANGCAT 720
Db 1248 AAGCCGGAAGAGGCAAGTCTCTGCTGCAAAAGGTGGCTCTGCNCGCATTAATTNANGCAT 1305
QY 721 TGGAGAACTGTGTTGAACACTCTCAAGATGGGGTTTCCCGGAAGCCGCACTTCC 780
Db 1306 TGGAGAACTGTGTTGAACACTCTCAAGATGGG--TTTCCGGAAGCCGCACTT-C 1362
QY 781 TTTTAGAATCATCAAAATTTGAAGCCNGTTCACCAAGCTCTGCAAGTATNATNGGTTAA 840
Db 1363 TTTTAGAATCATCAAAATTTGAAGCCNGTTCACCAAGCTCTGCAAGTATNATNGGTTAA 1419
QY 841 NGNAAAATCTNGTTTAAAGNGCATTTGAGAAGGAAACNTNAAAAGGGAANGGGGAACA 900
Db 1420 GGCAGAAATCTGTTTTTAA--GAGATTTGAGAGAAACTTGA---AGAGGAAGGGGAACA 1473
QY 901 ATTTGTGAAGAAAAATNCGNGGGAATTTTGGCCCTTCAAGGGGAANAATGGCCCTGG 960
Db 1474 GTTGTGAAGAAAAATCGGTGT----ATTTTGCCTTCAAGGTGAAG-ATGGCCCTGG 1527
QY 961 GGGGTAAAAGANGGCCACCTGGGG 985
Db 1528 GGGTAAAAGANGGCCACCTGGGTGGTG 1552

RESULT 5
ABI99762
ID ABI99762 standard; cDNA; 2152 BP.
XX AC ABI99762;
XX DT 07-MAR-2002 (first entry)
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:842.
XX KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX KM vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
OS Mus musculus.
XX PN WO200188188-A2.
XX PD 22-NOV-2001.

XX PF 18-MAY-2001; 2001WO-JP04192.
XX PR 18-MAY-2000; 2000JP-0145977.
XX PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX DR WPI; 2002-034733/04.
XX DR P-PSDB; ABB57301.
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
PS Claim 2; Page 2087-2091; 2690pp; English.
XX CC The present invention describes a method for examining ischaemic
XX CC conditions, comprising measuring the expression levels of particular
XX CC genes (I) in a test sample or determining the expression profile of a
XX CC gene group in the sample comprising genes selected from (I). The method
XX CC is useful for examining the ischaemic condition (e.g. compressive
XX CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX CC expression levels of particular genes (ABI99202 to ABI9912, encoding
XX CC the protein sequences in ABB57020 to ABB57374) or by determining the
XX CC expression profile of a gene group comprising these genes. The
XX CC expression levels or expression profiles produced by these genes are
XX CC used as an indicator when screening for ischaemic condition-improving
XX CC drugs or therapeutics for ischaemic diseases. ABI9913 and ABI9914
XX CC represent PCR primers for a mouse ischaemic condition related sequence,
XX CC which are used in the exemplification of the present invention.
SQ Sequence 2152 BP; 586 A; 455 C; 576 G; 535 T; 0 other;
Query Match 57.7%; Score 573.4; DB 24; Length 2152;
Best Local Similarity 78.9%; Pred. No. 1.6e-151;
Matches 765; Conservative 0; Mismatches 186; Indels 19; Gaps 7;
QY 1 CTEACCCAGTGTGCTCCTCAGATGTTGGGTATGCTGGAAGAAACATATGGAATAATG 60
Db 479 CACATCCGATTACTCCTCAGATGTTGGGTATGCTGGAAGAAACATATGGAATAATG 538
QY 61 GAACAAAATTTGAACACTTTGCCAAAATTTGGATGGAATAATCATTAACATTCACTTAATA 120
Db 539 GAACAAAATTTGAACACTTTGCCAAAATTTGGATGGAATAATCATTAACATTCACTTAATA 598
QY 121 ACCCGTATTTCCAGTTCAGAGATGATACAGTTTGAATGAAGTATGATGATTAAGAG 180
Db 599 ACAGTATTTCCAGTTCAGAGATGATACAGTTTGAATGAAGTATGATGATTAAGAG 658
QY 181 TTTTGGATTTTGGATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 240
Db 659 TTTTGGATTTTGGATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 718
QY 241 TTTTGGCCAGTGAAGCATTTGTCAGAGAAGTATGCGCTGCAATCCAAAGCTGTGAAATTT 300
Db 719 TTTTGGCCAGTGAAGCATTTGTCAGAGAAGTATGCGCTGCAATCCAAAGCTGTGAAATTT 778
QY 301 TGGCACAAGAATGATGATGATTTTGGCCAAGCTCGTTTGAAGAAAAAGCATTTATTA 360
Db 779 TGGCCACAGAGATGATGATGATTTTGGCCAAGCTCGTTTGAAGAAAAAGCATTTATTA 838
QY 361 TGGTTGGCTTTGATATGATTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGA 420
Db 839 TGGTTGGCTTTGATATGATTAAGAAGCTGCAGAAAAATGCTATGAGAAATCTGGCCTGA 898
QY 421 CACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCTTA 480
Db 899 CACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTAACCAAGCACTCTTA 958
QY 481 CTTATGAAGCACTGGACTCTGTCCAGAAAGCAAGGTGCAACGCTGTTGATAGAGAG 540

Db 959 CTTACGAGCACTGGGGCTCTGTCCAGAAAGACAAGGTGGAACCTGTGTGACAGAGGGG 1018
Qy 541 ATAAATCATATGAGGAAAGTGGGTCAATAATCCTAGTGTGACTGATTCCAAGGGAC 600
Db 1019 ACAACACTTACGAGGAAAGTGGGTCAATCAACCTAGTGAGGCTCATCTCAAAAGGAC 1078
Qy 601 ACCCACTAGGCGCTACAGTCTTGTCTAGTGTGCAGAACTGTGCTGGCAGCTGAGAGGG 660
Db 1079 ACCCACTGGGCGCAACAGTCTGTGCTAGTGCGCGAGCTGTGCTGGCAGCTGAGAGGG 1138
Qy 661 AAGCCGGAAGAGAGCAAGTTCCTGTGTGCAAGGTGGCTCTGCNGCATATTANGCAT 720
Db 1139 AAGCCGGAAG--GAGGAGGTTCCCGGGCAAGGTGGCTCTGCAGACAATCTAGGCT 1196
Qy 721 TGGAGAACTGTGTTGTGTACACTCTACACAGATGGGGTTTCCCGGAAGCCCGCAGTTCC 780
Db 1197 CGGAGGAGCTGTGTTGTCAACCTCTACAGGATGGG--TTTTCGGAAGCTGCCAGCTCC 1254
Qy 781 TTTTAGAACTCATCAAAATTTGAAGCCNGTTCACCAAGCTCTGCAGTATNGGTTAA 840
Db 1255 TTCAGGA--CGCACAGGTTTTCAGTGTCTCCCA--CCAGCTCTGCAGGGGATGATTCAA 1310
Qy 841 NGNAAAATCTNGTTTAAAGNGATGAGAAGAAACNTNAAAGAGGAANGGGAACA 900
Db 1311 GGCAAACTCGTCTTTAA--GGAGATTGAGAAGAACTTGA---AGAGGAAGGGAACA 1364
Qy 901 ATTTGGAAGAAAGAAATATGCGNGGGAATTTTGCCCTTCAAGGGGAANAATGGCCCTGG 960
Db 1365 GTTCGTGAAGAAATCGGTGGC-----ATTTTGCCCTTCAAGTGAAGATGGCCCTGA 1419
Qy 961 GGGGTAAAG 970
Db 1420 GGCAAGAAAG 1429

RESULT 6

ABX40513
ID ABX40513 standard; cDNA; 432 BP.
XX
AC ABX40513;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #5678.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle

XX
PS Claim 2; SEQ ID No 5678; 245bp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139.
XX
SQ Sequence 432 BP; 127 A; 75 C; 122 G; 108 T; 0 other;

Query Match 32.5%; Score 322.8; DB 25; Length 432;
Best Local Similarity 88.7%; Pred. No. 5.1e-81;
Matches 360; Conservative 0; Mismatches 44; Indels 2; Gaps 1;

Qy 357 AAAATGGTTGGCTTGATATGAGTAAAGAGCTGCAGAAATGCTATGAGAAATCTGGC 416
Db 2 AAAATGGTTGGTTTGATATGAGTAAAGAGCTGCCAGAAAGTGTATGAGAAATCTGGC 61
Qy 417 CTGACACCAATGATATGACGTAATAGAACTTCACGATGCTTTCTACCAACGAATC 476
Db 62 CTGAGACCAAGTATGACGTAATAGAACTTCACGATGCTTTCTGCCAATGAATCTT 121
Qy 477 CTTACTTATGAAGCACTGGGACTGTGCCAGAGCAAGGTGCAACGCTGTTGATAGA 536
Db 122 ATTACTTACGAAGCACTGGGACTGTGCCAGAGGTGCAAGAACTGTTGAAGA 181
Qy 537 GGAGATAATACATATGAGGAAAGTGGGTATTAATCCTAGTGTGACTGATTTCGAAG 596
Db 182 GGAGATAATACTTATGAGGAAAGTGGGTATTAATCCTAGTGGCGATTATTTCAAG 241
Qy 597 GGACACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAAACTCTGTGGCAGCTGAGA 656
Db 242 GGGCACCCACTTGAGGCTACAGGCTGTGGCTCAGTGCCTTGAACCTGTGCGCATCTGAGA 301
Qy 657 GGGGAAGCCGGAAGAGGCAAGTCTCTGTGTGCAAGAGTGGCTCTGCNGCATATTAN 716
Db 302 GGGGAAGCCGGAAG--AAGGCAAGTCTCTGTGTGCAAGAGTGTCTGTGAGCATATATAG 359
Qy 717 GCATTGAGGAAGTGTGTTGTAACACTCTACAAGATGGGGTTTTC 762
Db 360 GCATCGAGGAGAGCTGTGTTGTTACGCTGTACAAGATGGGCTTTCC 405

RESULT 7

ABL16479
ID ABL16479 standard; DNA; 1837 BP.
XX
AC ABL16479;
XX
DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 910.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX PS Claim 1; SEQ ID NO 910; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1837 BP; 428 A; 464 C; 528 G; 417 T; 0 other;

Query Match 26.8%; Score 266; DB 23; Length 1837;
Best Local Similarity 61.0%; Pred. No. 1.1e-64;
Matches 465; Conservative 0; Mismatches 292; Indels 5; Gaps 2;

QY 1 CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATG 60
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
553 CTGGCCCATGGCTGCCAGATCTTCGGCAATGCGGCAAGAGACATGAAGAATATG 612

QY 61 GAACAAAAATGAACACTTTCGAAAAATGGATGAAAAATCATAAACATTCACTTAATA 120
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
613 GCACTAAGCCCCGAGCATTTCCGCCAAGATCGCTTGAAGAACCACAGCACTCCGTCATA 672

QY 121 ACCCGTATTCAGTCCAGATGAAATACAGTTAGATGAAGTATGGCATCTAAGAAG 180
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
673 ATCCCTACTCGCAGTTCGCGGATGAAATACACCTGGAGCAGATTATGAAGTCGCCCAGG 732

QY 181 TTTTGA---TTTGTGACTATCTTAACAATGTGTGCCCACTTCAGATGTGTGCAGCAG 237
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
733 TTGTGAGGAGAGTGTGACCAAGCTGCAGTGTGTCCCACTTCGATGATCTGAGCTG 792

QY 238 CAATTTGGCCAGTGAAGCATTTGTACAGAAGTATGCTGCAATCCAAAGCTGTGAAA 297
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
793 CCAATTCTCGCTCCGAGGCTTCGTGCGTCGCCACGAGATTGAAAAAGAGCTGTGAAA 852

QY 298 TTTTGGCACAAGAATGATGACTGTTTGGCAAGCTCGTTTGAAGAAAAAGCATTATTA 357
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
853 TTGTAGGCATGAGATGGCCAGTACCCGGCGTCCACCTTTGCCGACAAGAGCTTGATGA 912

QY 358 AAATGTTGGCTTGTATATGATGAAGAAGCTGCAAGAAAAATGCTATGAGAAATCTGGCC 417
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 913 AGATTGCTGAACCGATATGACCCGTCCTGGCCACCACGCGTCTGTTCGAAAGATGGGT 972
QY 418 TGACACCAATGATATATGACGTAATAGAACTTCACGATGCTTTTCTACCAAGAACTCC 477
Db | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
973 ACAAGCCCGAGATGTCAGGTTGTGAACTGCACAGATGCTCTCGGCCAACGAGTTGA 1032

QY 478 TTACTTATGAAGCACTGGACTCTGTCCAGAAGACACAAGTGCAACGCTGGTTGATAGAG 537
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1033 TCACGTATGAGGCACTTGACTGTGTGAGAGGGCAACGCCGGCAGTTTCATCGACGCTG 1092

QY 538 GAGATATACATATGAGGAAAGTGGTCATTAATCCTAGTGGTGAATTTCAAAGG 597
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1093 GAGACAAACACTACGCTGGCAAGTTCTGTGTCAAACCCAGTGTGTCTGATCTCGAAGG 1152

QY 598 GACACCCACTAGGCGCTACAGTCTTCTCAGTGTGCAAGACTCTGTCGACGCTGAGAG 657
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1153 GCCATCCTCTGGGCGCCACGGGCTTGGCACAATGTGCTGAGCTCTGCTGGCAGCTCCGTG 1212

QY 658 GGAAGCCGGAAGAGGCAAGTTCCTGTGCAAGGTGCTTGCGCATTAATTANG 717
Db | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1213 GATTGGCTG--AGAAGCGCCAGGTGCCCAATGCTCAGTTGGCTTTGCAGACAAATCTGGG 1270

QY 718 CATTGAGGAATGTTGTTGTAACTCTACAGATGGGTT 759
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1271 ACTGGTGTGCGCTGTGTGTTGCTTGTATCGTCTGGGTTT 1312

RESULT 8
ABL16661
ID ABL16661 standard; DNA; 1237 BP.
XX ABL16661;
AC 26-MAR-2002 (first entry)
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 1456.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX PS Claim 1; SEQ ID NO 1456; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1237 BP; 289 A; 315 C; 369 G; 264 T; 0 other;

Query Match 26.2%; Score 260; DB 23; Length 1237;
Best Local Similarity 60.7%; Pred. No. 4.6e-63;
Matches 459; Conservative 0; Mismatches 292; Indels 5; Gaps 2;

QY 7 CAGTTGCTCTCAGATGTTGGGTATGCTGAAAAAGACATATGAAAAATATGGAACA 66
DB 394 CATTGTCCGCAAGATCTTCGGCAACGCCGCAAGAGACACATGAAGAATATGGCACTA 453
QY 67 AAATTGAACACTTTCAAAAAATTGATGAAAAATCATAAACATTCAATTAAACCCGT 126
DB 454 AGCCCGAGATTTCGGCAAGATCGCTTGAAAGAACACAAAGCACTCCGTCAATAATCCCT 513
QY 127 ATTCCAGTTCCAAGATGAATACAGTTAGATGAAGTGAAGCACTAAAGAGTTTGG 186
DB 514 ACTCGAGTTCGCGATGAATACACCCGAGACAGATTATGAAGTCGCCCCAGTTGTGG 573
QY 187 A---TTTTGACTATCTTACAATGTGTCCCACTTCAGATGTGCTGCAGCAATT 243
DB 574 AGGAGTGTGACCAAGCTGCAGTGTCTCCCACTTCGATGATCTGGAAGCTGCCATT 633
QY 244 TGGCCAGTGAAGCATTGTACAGAGATGAGCCCTGCAATCCAAAGCTGGAATTTGG 303
DB 634 TCGCCTCCGAGGCTTCGTGCGTCCGACGAGTTGAAAAGCAGGCTGTGAGATTGTGG 693
QY 304 CACAAGAAATGATGACTGATTGTCGAAGCTCGTTGAAAGAAAAAGCATTATTAATG 363
DB 694 GCATGAGATGGCCAGTGAACCCGCGCTCCACTTTGCCAGACAGAGCTGATGAAGATT 753
QY 364 TTGGCTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGAC 423
DB 754 CTGGAACCATATGACCCGCTGTGGCCACGACGCTGTTCGCAAGAGTGGTACAAGC 813
QY 424 CAAATGATATGACGTAATAGAACTTCACGATTGCTTTCTACCAAGAACTCCTACTT 483
DB 814 CCCAGATGTCCAGGTGTGGAACGTGACGATGCTTCGCGCAAGAGTGTATCAGT 873
QY 484 ATGAAGCACTGGGACTCTGTCCAGAAGCAAGGTGCAACGCTGTTGATAGAGAGATA 543
DB 874 ATGAGGCACTTGGACTGTGTGAGAGGGCAAGGCCGCGAGTTCATCGACGCTGAGACA 933
QY 544 ATACATATGAGGAAAGTGGTCAATAATCCTAGTGTGAGTGAATTTCAAAAGGACACC 603
DB 934 ACACCTACGCTGGCAAGTTCGTGTCAACCCAGTGTGTGTCTGATCTCGAAGGGCCATC 993
QY 604 CACTAGGCGCTACAGGTCTTGTCTCAGTGTGCAAGACTCTGTGCGAGCTGAGAGGGGA 663
DB 994 CTCTGGGCGCACGGGTCTGGCACAAATGTGCTGAGCTCTGTGCAAGCTCCGTGATTGG 1053
QY 664 CCGGAAAGAGGCAAAAGTTCCTGCTGCAAGAGTGGCTCTGCGCATTAATTANGCATGG 723
DB 1054 CTG--AGAGGCGCAGGTGCCCAATGCTCAGTTGGCTCTGCAGCACAATCTGGACTGG 1111
QY 724 AGGAAGTGTGTTGTATACACTCTACAAGATGGGGTT 759
DB 1112 TGGTGTGTGTGTGTGCTTGTATCTGTGGGTTT 1147

RESULT 9
ABX42390
ID ABX42390 standard; cDNA; 346 BP.
XX
AC ABX42390;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #7555.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle -
XX
PS Claim 2; SEQ ID No 7555; 245bp; English.

XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139..
XX

SQ Sequence 346 BP; 96 A; 66 C; 99 G; 85 T; 0 other;

Query Match 25.5%; Score 253; DB 25; Length 346;
Best Local Similarity 86.8%; Pred. No. 2.5e-61;
Matches 289; Conservative 0; Mismatches 42; Indels 2; Gaps 1;

QY 430 ATATTGACGTAATAGAACTTCAGATGCTTTTCTACCAAGAACTCCTTACTTATGAAG 489
DB 14 ATCCGACATATAAGAACTTCAGATGCTTTTCTGCCAATGAACCTTATTACTTACGAAG 73
QY 490 CACTGGAGCTCTGTCCAGAAGACAAGGTGCAACGCTGTTGATAGAGAGATAATACAT 549
DB 74 CACTGGAGCTGTGTCCAGAAGGTCAAGGTGAAAACCTGTTGAAAAGAGAGATAATACAT 133
QY 550 ATGAGAGAAAGTGGGTGATAAATCCTAGTGTGAGTGAATTTCAAGGAGACACCACTAG 609
DB 134 ATGAGAGAAAGTGGGTGATAAATCCTAGTGTGAGTGAATTTCAAGGAGACCACTTG 193

```

Oy      610 GCGGTACAGGCTTGTGCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGGGAAAGCCGGAA 669
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      194 GAGGTACAGGCTGTGCTCAGTGTGCGTTGAACTCTGCTGGCATCTGAGAGGGGAAAGCCGGAA 253
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      670 AAGAGCAAGTTCCTGTGTGCAAAAGGTGCTCTGCNGCATATATTANGCATTTGAGGAAC 729
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      254 A--AAGCAAGTTCCTGTGTGCAAAAGGTGCTCTGCAGCATATATATAGGCATCGGATGAAC 311
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy      730 TGTGCTGTTAACACTCTACAAGATGGGGTTTC 762
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      312 TGTGCTGTTAACGCTGTACAGATGGGCTTCC 344
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
ABX35482
ID      ABX35482 standard; cDNA; 314 BP.
XX
XX      ABX35482;
AC
XX
XX      20-FEB-2003 (first entry)
DT
XX
XX      Bovine EST associated with lactation/muscle/fat deposition #647.
DE
XX
XX      Bovine; BF; EST; expressed sequence tag; lactation; LMFD;
KM      muscle deposition; fat deposition; genome mapping; gene identification;
KM      gene analysis; cattle breeding.
XX
XX      Bos Taurus.
OS
XX
XX      US2002137139-A1.
PN
XX
XX      26-SEP-2002.
PD
XX
XX      24-SEP-2001; 2001US-0960352.
PF
XX
XX      12-JAN-1999; 99US-115707P.
PR      11-JAN-2000; 2000US-0480902.
XX
XX      (BYAT/) BYATT J C.
PA      (MATH/) MATHIALAGAN N.
PA      (TAON/) TAO N.
PA      (WARR/) WARREN W C.
XX
XX      Byatt JC, Mathialagan N, Tao N, Warren WC;
PI
XX
XX      WPI; 2003-110599/10.
XX
XX      New nucleic acid associated with lactation, and muscle and fat
PT      deposition, useful for genome mapping, gene identification and
PT      analysis, cattle breeding, or for genetically improving cattle -
PT
XX
XX      Claim 2; SEQ ID No 647; 245pp; English.
XX
XX      The invention relates to a purified nucleic acid molecule associated with
CC      lactation or muscle and fat deposition (designated LMFD), derived
CC      from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC      second nucleic acid molecule comprising any of 15112 nucleotide
CC      sequences, appearing as ABX34836-ABX4947, or complements of them.
CC      Also included are; (1) a transformed cell having a nucleic acid
CC      comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC      translated sequence that functions in the cell to cause termination of
CC      transcription and addition of polyadenylated ribonucleotides to a 3' end
CC      of the mRNA molecule; and (2) determining a level or pattern of a
CC      molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC      nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC      complement or fragment) with a complementary nucleic acid molecule
CC      obtained from the bovine cell or tissue, where hybridisation between the
CC      marker nucleic acid and the complementary nucleic acid permits the
CC      detection of the molecule; and (b) detecting the level or pattern of the
CC      complementary nucleic acid, where the detection of the complementary
CC      nucleic acid is predictive of the level or pattern of the molecule.
CC      The LMFD nucleic acid is used for determining a level or pattern

```

CC	of a molecule in a bovine cell or tissue. It is useful for genome
CC	mapping, gene identification and analysis, cattle breeding, preparation
CC	of constructs for use in cattle gene expression, or for genetically
CC	improving cattle. The present sequence is one of the 15112 bovine
CC	LMFD EST (expressed sequence tag) nucleic acids.
CC	Note: The present sequence was not shown in the specification but
CC	was obtained in electronic format from the USPTO web site:
CC	seqdata.uspto.gov/sequence.html?DocID=20020137139.
XX	
SQ	Sequence 314 BP; 95 A; 51 C; 91 G; 77 T; 0 other;
	Query Match 23.9%; Score 238; DB 25; Length 314;
	Best Local Similarity 86.9%; Pred. No. 4.1e-57;
	Matches 273; Conservative 0; Mismatches 40; Indels 1; Gaps 1
OY	358 AAATGTTGGCTTGTGATAGTAAGAAAGCTGCAGAAGAAATGCTATGAGAAATCTGGCC 417
Db	1 AAATGTTGGTTTGATATGAGTAAAGAAAGCTGCCAAGGTGCTATGAGAAATCTGGCC 60
OY	418 TGACACCAATGATATTGACGTAATAGAACTTCACGATTGCTTTCTACCACGAAGTCCC 477
Db	61 TGAGACCAAGTGATATTGACGTAATAGAACTTCACGAACGCTTGCGCCAATGAAGTTA 120
OY	478 TTACTTATGAAGCAGCTGGGACTCTGTCCAGAAAGACAAGGTGCACGCTGGTTAGAG 537
Db	121 TTACTTACGAAGCAGCTGGGACTGTGTCCAGAAATGTCAAAGGTGAAAACTGGTTGAAAGAT 180
OY	538 GAGATAATACATATGAGGAAAAGTGGGTCAATTAATCCTAGTGGTGAGCTGATTTCAAAGG 597
Db	181 GAGATAATACTTATGAGGAAAAGTGGGTCAATTAATCCTAGTGGCGGATTAATTTCAAAGG 240
OY	598 GACACCCAGTAGCGCTACAGGCTTGCTCAGTGTGCAGAACTGCTGGCAGC-TGAGA 656
Db	241 GGCAACCCAGCTTGAGAGCTACAGGCTGTGGCTCAGTGCCTGGAAGTCTGCTGGCATCTTGAGA 300
OY	657 GGGGAAAGCCGGAAA 670
Db	301 CGGGGAAGCGGAAA 314
RESULT 11	
AAI99682/c	
ID	AAI99682 standard; DNA; 4411529 BP.
XX	
AC	AAI99682;
XX	
DT	15-JAN-2002 (first entry)
XX	
DE	Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
XX	
KW	Mycobacterium tuberculosis; strain H37Rv; strain CDC 1551; genome;
KW	variation; epidemiology; patient treatment; epidemic monitoring; ds.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	US6294328-B1.
XX	
PD	25-SEP-2001.
XX	
Pf	24-JUN-1998; 98US-0103840.
XX	
PR	24-JUN-1998; 98US-0103840.
XX	
PA	(GENO-) INST GENOMIC RES.
XX	
PI	Fleischmann RD, White OR, Fraser CM, Venter JC;
XX	
DR	WPI; 2001-647261/74.
XX	
PT	Evaluating strain variation of Mycobacterium tuberculosis, comprises
PT	determining the nucleotide sequence of the strain at positions in the
PT	genome corresponding to positions where M. tuberculosis strains CDC
PT	1551 and H37Rv differ -

XX Claim 3; SEQ ID NO 1; 3pp + Sequence Listing; English.
PS
XX
CC The invention relates to evaluating strain variation within and between
CC different populations of the tuberculosis bacterial pathogen,
CC Mycobacterium tuberculosis or related Mycobacterium by determining the
CC nucleotide sequence of the first strain at positions in the complete
CC sequence of the genome that correspond to positions that differ in the
CC nucleotide sequences of M. tuberculosis strains CDC 1551 (A199683) and
CC H37Rv (A199682). The method is useful for evaluating strain variation of
CC M. tuberculosis and has valuable application in the fields of
CC tuberculosis genetics, epidemiology, patient treatment and epidemic
CC monitoring.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=6294328B1.
XX
XX Sequence 441529 BP; 758565 A; 1443983 C; 1444602 G; 758379 T; 0 other;

Query Match 21.5%; Score 214; DB 22; Length 441529;
Best Local Similarity 57.1%; Pred. No. 1.9e-48;
Matches 428; Conservative 0; Mismatches 317; Indels 5; Gaps 2;

OY 6 CCAGTTGCTCCTCAGATGTTGGTATGCTGAAAAAGACATATGAAAAATATGAACA 65
Db 3099716 CCGTGCGCGCGTGATGTTGGGGCGCGCGCGCGAACAACATGAAGAAATACGCACCC 3099657
OY 66 AAAATGACACTTTTGCAAAAATTTGATGAAAAATCATAAACATTCAGTTAATACCCG 125
Db 3099656 ACCGCGAGCATTTCCGGAAGATCGGCTACAAGAACCAAGACACTCGGTCAACAACCCG 3099597
OY 126 TATTCAGTTCCAGATGAATACAGTTTAGATGAAGTATGCACTTAAGAAGTTT 185
Db 3099596 TATGACAGTTTCAGAGACGAATACACCTTGAGCAGACATCTTGCCCTCAAAGATGATTTCC 3099537
OY 186 GATTTTTCATCTTACATGTTGTTCCCACTTCAGATGTTGCTGACAGCAATTTTG 245
Db 3099536 GACCCGCTGACCAAAATTGCAAGTCTCTCCACCTCCGACGGGTGCGCGCGGTGCTG 3099477
OY 246 GCCAGTGAAGCATTTGTACAGATATGGCTGCAATCCAAAGCTGTGAAATTTTGCA 305
Db 3099476 GCCAGTGAAGATTAACCTGCGCAACCAACCTTGCGCGCGGTGGAATCTGCGG 3099417
OY 306 CAAAGATGATGATGATTTGCAAGCTGTTTGAAGAAAAAGCATTTAAATGTT 365
Db 3099416 CAGCGATGACCAACCGACTTGTCTTCCACTTTGATG--GCAGTGCCCGCAATATCATC 3099360
OY 366 GCGTTGATATGAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCTGACACCA 425
Db 3099359 GCGTACGACATGACTGTGCAAGCGGCAACAGGTTTACAGCAATCCGGACTCGGCCG 3099300
OY 426 AATGATATGACGTAATAGAACTTCAAGATTGCTTTCTTACCAAGAACTCCTTACTTAT 485
Db 3099299 AAGGACTTCGAGATGATCGAGCTGCAGACTGCTTCTCAGCCAAAGACTACTGCTTAC 3099240
OY 486 GAAGCACTGGGACTCTGTCCAGAGCAAGGTGCAACGCTGTTGATAGAGAGATTAAT 545
Db 3099239 GAGCGCTCGGCTGTGCGGCGCGGCGCGCGCGAGCTGATGACGACAACAGAGACC 3099180
OY 546 ACATATGAGGAAAAAGTGGGTATTAATCTAGTGTGAGTGAATTTCAAGGACACCCA 605
Db 3099179 ACCTACGCGGAGCGCTGGGTCTGCAACCATCCGGTGGGTGATCTCCAAGGCCATCCG 3099120
OY 606 CTAGCGCTACAGGCTCTGCTCAGTGTGAGAACTCTGCTGCGAGCTGAGAGGGAAGCC 665
Db 3099119 CTGGGTGCGACGGGGGTGCGCAGTGGCGGAGCTGACCTGGCAGCTGCGCGCACCGCC 3099060
OY 666 GGAAGAGAGCAAAAGTTCCTGTGCAAAAGGTGGCTCTGCNGCATTAATTANGCATTTGAG 725
Db 3099059 G--AGCGCGCGCAGGTGACAAACGTGACCGCGCTGCAACACATCGGGCTGGGCG 3099002
OY 726 GAAGTGTGTTGTAACACTCTCAAGATGG 755

Db 3099001 GCGCGCGCGGTGTACCGCATACCAACGGG 3098972
RESULT 12
ABL16660/c
ID ABL16660 standard; DNA; 3613 BP.
XX
XX ABL16660;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1453.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PE
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Claim 1; SEQ ID NO 1453; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3613 BP; 948 A; 876 C; 843 G; 946 T; 0 other;

Query Match 21.3%; Score 211.4; DB 23; Length 3613;
Best Local Similarity 60.3%; Pred. No. 4.1e-49;
Matches 384; Conservative 0; Mismatches 248; Indels 5; Gaps 2;

OY 126 TATTCAGTTCCAGATGAATACAGTTTAGATGAAGTATGCGATCTTAAGAGTTT 185
Db 1725 TACTCGAGTTCGCGATGAATACACCTGAGAGATTAAGTGGCCCCAGGTGTG 1666
OY 186 GA--TTTTCAGTAATCTTACATGTTGTCCCACTTCAGATGTTGCTGCAGCAAT 242
Db 1665 GAGGAGTGTGACCAAGCTGACGTGTCTCCCACTTCGATGATGTGAGCTGCCATT 1606
OY 243 TTGGCCAGTGAAGCATTTGTACAGAGTATGGCTGCATCCAAAGCTGTGAAATTTG 302
Db 1605 CTCGCCCTCCGAGGCTTGTGCTGCGCCACGGAATTGAAAAAGCAGGCTGTGAGATTGTG 1546
OY 303 GCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTTAAATG 362
Db 1545 GGCATGAGATGGCAGTGAACCGCGGCTCCACTTTTGCCGACAAGAGCTTGATGAAGATT 1486

XX 23-MAR-2001; 2001WO-US09231.
PF 23-MAR-2000; 2000US-191637P.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 31951; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4082 BP; 1085 A; 989 C; 929 G; 1079 T; 0 other;

Query Match 21.3%; Score 211.4; DB 23; Length 4082;
Best Local Similarity 60.3%; Pred. No. 4.3e-49;
Matches 384; Conservative 0; Mismatches 248; Indels 5; Gaps 2;
QY 126 TATGCCAGTTCAGATGAATAGATTAGATGAAGTGCATCTAAGAAGTTT 185
Db 3953 TACTCGAGTTCGCGATGATACACCTGAGACAGATTATGAGTCGCCAGTTGTG 3894
QY 186 GA---TTTGTGACTATCTTACAATGTTGCCACTTCAGATGCTGCAGCAAT 242
Db 3893 GAGGAGTGTGACCAAGCTGCAGTGTCTCCACTCCGATGATCTGAGCTGCCATT 3834
QY 243 TTGCCAGTGAAGCATTGTGTACAGAGATGCGCTGCAATCCAAGCTGTGAAATTTG 302
Db 3833 CTGCTCCGAGGCCCTTCGTGCTGCCACGAGATTGAAAGCAGCTGTGAGATTGTG 3774
QY 303 GCACAGAAATGATGACTGATTGCCAAGCTGTTGAAGAAAAAGCATTATTAATG 362
Db 3773 GGCATGAGATGCGCAGTGAACCGCGCTCCACTTGGCCAGACAGCTTGATGAAGATT 3714
QY 363 GTTGCTTTGATGATGAATAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCTGACA 422
Db 3713 GCTGAACCGATATGACCCGCTGCGCACCGAGCGTCTGTCGCCAAGAGTGGTACAAG 3654
QY 423 CCAATGATATGACGTAATAGACTTCAGATTTGCTTTCTACCAACGAATCCTTACT 482
Db 3653 CCCCAGATGTCCAGGTGTGGAATGACGATTGCTTCGGCCAAAGATGATCAGC 3594
QY 483 TATGAAGCACTGGGACTCTGTCCAGAAAGACAAAGTGCACCGCTGTTGATAGAGAGAT 542
Db 3593 TATGAGGCACTTGACTGTGTGAGAGGGGCAAGCGCGGAGTTTCAATGACGCTGAGAC 3534
QY 543 AATACATATGAGGAAAGTGGGTCTAATAATCTAGTGTGAGTGAATTTCAAGGGACAC 602
Db 3533 AACACCTACGCTGCGCAAGTTCGTGTCAACCCAGTGTGTTGATCTCGAAGGGCCAT 3474
QY 603 CCACTAGCGGCTACAGTCTTGTCTCAGTGTGCAAGACTCTGCTGCAGCTGAGAGGGAA 662
Db 3473 CCTCTGGGGCCACGGGTCTGACAAATGTGCTGAGCTCTGCTGCGAGCTCCGTGATTTG 3414
QY 663 GCCGAAAGAGGCAAGTTCCTGCTGCAAGAGTGCTCTGCGCATATTANGCATTTG 722

Db 3413 GCTG--AGAGCGCCAGGTGCCCAATGCTCAGTTGGCTCTGCAGACACAATCTGGGACTGG 3356
QY 723 GAGGAAGTGTGTTGTAACACTCTACAAGATGGGTT 759
Db 3355 GTGTGCTGTTGTGTTGCTTGTATGCTGTGGGTTT 3319
RESULT 15
ABL16478
ID ABL16478 standard; DNA; 4356 BP.
XX
AC ABL16478;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 907.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 907; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4356 BP; 1165 A; 942 C; 1021 G; 1228 T; 0 other;
Query Match 20.9%; Score 208.2; DB 23; Length 4356;
Best Local Similarity 60.0%; Pred. No. 3.5e-48;
Matches 382; Conservative 0; Mismatches 250; Indels 5; Gaps 2;
QY 126 TATGCCAGTTCAGATGAATACAGTTTAGATGATGATGCGCATCTAAGAAGTTT 185
Db 2137 TACTCGAGTTCGCGATGAATACACCTGAGACAGATTATGAAGTCCGCCAGTTGTG 2196
QY 186 GA---TTTGTGACTATCTTACAATGTTGCCACTTCAGATGCTGCAGCAGCAATT 242
Db 2197 GAGGAGTGTCTGACCAAGCTGCAAGTGTGTCTCCACTTCGATGATCTGAGCTGCATT 2256
QY 243 TTGCCAGTGAAGCATTGTGACAGAGATGCGCTGCAATCCAAGCTGTGAAATTTG 302
Db 2257 CTGCGCTCCGAGGCGCTTCGTGCTGCGCACGAGTGGAAAGCAGGCTGTGAAATTTGTA 2316

```
OY      303 GCACAAGAAATGATGACTGATTTCCCAAGCTCGTTTGAGAAAAAAGCATTTAAATG 362
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2317 GGCATGGAAGATGGCCAGTGAACCCGGCGTCCACCTTTGCCGACAAGAGCTTGATGAAGATT 2376
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      363 GTTGCTTTGATATGATGAATAAGAGCTGCAAGAAAAATGCTATGAGAAATCTGGCTGACA 422
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2377 GCTGAAACGATATGACCCGCTCTGGCCACCCAGCGCTCTGTTGCAAAAGAGTGGTACAAAG 2436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      423 CCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTCTAACCAACGAACCTCTTACT 482
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2437 CCCAAGGATGCCAGGTGTGGAACGACGATTGCTTCGGCCAAAGATTGATCACG 2496
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      483 TATGAAGCACTGGACTCTGTCCAGAAGACAGAGTGCAAGCTGTTGATAGAGAGAT 542
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2497 TATGAGGCACCTTGACTGTGTGAGAGAGGCAACGCCGGCGAGTTCATCGACGCTGGAGAC 2556
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      543 AATACATATGAGAGGAAAGTGGGTCAATAATCCTAGTGTGACTGATTTCAAAGGGACAC 602
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2557 AACACCTACGGTGGCAAGTTCGTGCTCAACCCCAAGTGTGTCTGATCTCGAAGGGCCAT 2616
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      603 CCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGGGAA 662
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2617 CCTCTGGGCGCCACGGGCTGTGCACAATGTGCTGAGCTCTGCTGGCAGCTCCGTGAATTG 2676
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      663 GCCGAAAAAGAGGCAAAAGTTCCTGCTGCAAAAGGTGCTCTGCTGCAATTAATTANGCAATTG 722
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2677 GCTG--AGAAAGCGCAGGTGCCAATGCTCAGTTGGCTTTGCAGACACAATCTGGGACTGG 2734
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY      723 GAGGAACCTGTGTTGTAACACTCTACAAGATGGGTT 759
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2735 GTGGTGGCCGTTGTGTTGCCCTTGATCGTCTGGGTTT 2771
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 27, 2003, 06:23:51
Job time : 334.626 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 06:06:02 ; Search time 374.822 Seconds
(without alignments)
8737.270 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 994
Sequence: 1 ctcaccacgtgtctctccag.....ccaccctggggtgtgggat 994

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	978	98.4	994	9	US-09-835-992A-21 Sequence 21, Appl
2	322.8	32.5	432	10	US-09-960-352-5678 Sequence 5678, Ap
3	253	25.5	346	10	US-09-960-352-7555 Sequence 7555, Ap
4	238	23.9	314	10	US-09-960-352-647 Sequence 647, App
5	196.2	19.7	1236	14	US-10-156-761-5843 Sequence 5843, Ap
6	196.2	19.7	9025608	14	US-10-156-761-1 Sequence 1, Appli
7	189.8	19.1	412	10	US-09-960-352-2946 Sequence 2946, Ap
8	92	9.3	249	10	US-09-878-574-8890 Sequence 8890, Ap
9	89.4	9.0	553	14	US-10-102-524-364 Sequence 364, App
10	67.2	6.8	938	12	US-10-027-632-34138 Sequence 34138, A
11	67.2	6.8	938	12	US-10-027-632-34139 Sequence 34139, A
12	67.2	6.8	938	13	US-10-027-632-34138 Sequence 34138, A
13	67.2	6.8	938	13	US-10-027-632-34139 Sequence 34139, A
14	42.6	4.3	2432	12	US-10-027-632-103694 Sequence 103694,
15	42.6	4.3	2432	12	US-10-027-632-103695 Sequence 103695,
16	42.6	4.3	2432	12	US-10-027-632-111751 Sequence 111751,

C	17	42.6	4.3	2432	13	US-10-027-632-103694	Sequence 103694,
C	18	42.6	4.3	2432	13	US-10-027-632-103695	Sequence 103695,
C	19	42.6	4.3	2432	13	US-10-027-632-111751	Sequence 111751,
C	20	40.2	4.0	3673778	12	US-10-312-841-2	Sequence 2, Appli
C	21	39.8	4.0	6313	8	US-08-781-986A-201	Sequence 201, App
C	22	39.4	4.0	671	14	US-10-184-644-346	Sequence 346, App
C	23	39.4	4.0	671	14	US-10-184-634-346	Sequence 346, App
C	24	39	3.9	580073	12	US-10-205-220-1	Sequence 1, Appli
C	25	38.6	3.9	9207	12	US-10-240-453-307	Sequence 307, App
C	26	38.2	3.8	1659	9	US-09-815-242-4442	Sequence 4442, Ap
C	27	38.2	3.8	1662	9	US-09-815-242-8331	Sequence 8331, Ap
C	28	38	3.8	1161	12	US-10-027-632-260949	Sequence 260949,
C	29	38	3.8	1161	13	US-10-027-632-260949	Sequence 260949,
C	30	38	3.8	10279	12	US-10-311-455-1563	Sequence 1563, Ap
C	31	37.6	3.8	3673778	12	US-10-312-841-1	Sequence 1, Appli
C	32	37.4	3.8	494	12	US-10-029-386-20259	Sequence 20259, A
C	33	37.4	3.8	1269	10	US-09-881-752A-91	Sequence 91, Appl
C	34	37.2	3.7	1161	12	US-10-027-632-260946	Sequence 260946,
C	35	37.2	3.7	1161	12	US-10-027-632-260947	Sequence 260947,
C	36	37.2	3.7	1161	12	US-10-027-632-260948	Sequence 260948,
C	37	37.2	3.7	1161	13	US-10-027-632-260946	Sequence 260946,
C	38	37.2	3.7	1161	13	US-10-027-632-260947	Sequence 260947,
C	39	37.2	3.7	1161	13	US-10-027-632-260948	Sequence 260948,
C	40	37.2	3.7	640681	10	US-09-790-988-1	Sequence 1, Appli
C	41	37	3.7	810	12	US-10-027-632-159529	Sequence 159529,
C	42	37	3.7	810	12	US-10-027-632-159530	Sequence 159530,
C	43	37	3.7	810	13	US-10-027-632-159529	Sequence 159529,
C	44	37	3.7	810	13	US-10-027-632-159530	Sequence 159530,
C	45	37	3.7	7589	12	US-10-240-453-264	Sequence 264, App

ALIGNMENTS

RESULT 1
US-09-835-992A-21
; Sequence 21, Application US/09835992A
; Patent No. US20020037541A1
; GENERAL INFORMATION:
; APPLICANT: Odate, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AN
; FILE REFERENCE: L0461/7112
; CURRENT APPLICATION NUMBER: US/09/835, 992A
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 08/896,164
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 994
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (705)..(705)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (716)..(716)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (807)..(807)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (830)..(830)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (833)..(833)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (841)..(841)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure

```

; LOCATION: (843)..(843)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (851)..(851)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (862)..(862)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (879)..(879)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (881)..(881)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (892)..(892)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (918)..(918)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (922)..(922)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (949)..(949)
; OTHER INFORMATION: n = a, c, g or t
; NAME/KEY: Unsure
; LOCATION: (972)..(972)
; OTHER INFORMATION: n = a, c, g or t
;
US-09-835-992A-21

```

Query Match	98.4%;	Score 978;	DB 9;	Length 994;
Best Local Similarity	100.0%;	Pred. No. 2.9e-248;		
Matches 994;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY		1	CTCACCAGTGTGCTCCCTCAGATGTTGGGTATGCTGAAAAAACAATATGAAAAATATG	60
Dp		1	CTCACCCAGTGTCTCCTCAGATGTTGGGTATGCTGAAAAAGAACATATGAAAAATATG	60
OY		61	GAACAAAATTGAACACTTTGCCAAAAATTGGATGAAAAATCATAAACATTCACTTAATA	120
Dp		61	GAACAAAATTGAACACTTTGCCAAAAATTGGATGAAAAATCATAAACATTCACTTAATA	120
OY		121	ACCCGATATCCCAGTTCACAAGATGAATACAGTTTAGATGAAGTGATGGCATCTAAGAAG	180
Dp		121	ACCCGATATCCCAGTTCACAAGATGAATACAGTTTAGATGAAGTGATGGCATCTAAGAAG	180
OY		181	TTTTTGATTTTGTACTATCTTACATGTTGTCCCACTTCAGATGCTGTGCAGCAGCAA	240
Dp		181	TTTTTGATTTTGTACTATCTTACATGTTGTCCCACTTCAGATGCTGTGCAGCAGCAA	240
OY		241	TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCCTGCAATCCAAGCTGTGAAATTT	300
Dp		241	TTTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCCTGCAATCCAAGCTGTGAAATTT	300
OY		301	TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTAATAAA	360
Dp		301	TGGCACAAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTAATAAA	360
OY		361	TGGTTGGCTTTGATATGAGTAAGAAGCTGCCAAGAAATGCTATGAGAAATCTGGCCCTGA	420
Dp		361	TGGTTGGCTTTGATATGAGTAAGAAGCTGCCAAGAAATGCTATGAGAAATCTGGCCCTGA	420
OY		421	CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAATCCTTTA	480
Dp		421	CACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAATCCTTTA	480
OY		481	CTTATGAAGCACTGGGACTCTGTCCAGAAGACAAGGTGCAACGCTGTTGATAGAGAG	540
Dp		481	CTTATGAAGCACTGGGACTCTGTCCAGAAGACAAGGTGCAACGCTGTTGATAGAGAG	540
OY		541	ATAATACATATGAGGAAGTGGGTCAATAATCCTAGTGTGACTGATTTCAAAAGGAC	600

Db	541	ATAATACATATGGAGGAAGTGGGTCAATAATCCTAGTGGTGA	CTGATTTCAAAAGGCAC	600
QY	601	ACCCACTAGGCGCTACAGGCTCTTGCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGGG		660
Db	601	ACCCACTAGGCGCTACAGGCTCTTGCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGAGGGG		660
QY	661	AAGCCGAAAAAGAGGCAAAAGTTCCGTGTGCAAAAGGTGGCTCTGCNGCATTAATTNANGCAT		720
Db	661	AAGCCGAAAAAGAGGCAAAAGTTCCGTGTGCAAAAGGTGGCTCTGCNGCATTAATTNANGCAT		720
QY	721	TGGAGGAACCTGTGTTGTATACTCTACAAGATGGGGTTTTCCCGGAAGCCGACAGTTCC		780
Db	721	TGGAGGAACCTGTGTTGTATACTCTACAAGATGGGGTTTTCCCGGAAGCCGACAGTTCC		780
QY	781	TTTTAGAACTCATCAAATTTGAAGCCNGTTCCAACCAAGCTCTGCAAGTNAATNGGTTAA		840
Db	781	TTTTAGAACTCATCAAATTTGAAGCCNGTTCCAACCAAGCTCTGCAAGTNAATNGGTTAA		840
QY	841	NGNAAAATCTNGTTTTAAAGNGGATTTGAGAAGGAACNTNAAAGAGGGGANGGGGAACA		900
Db	841	NGNAAAATCTNGTTTTAAAGNGGATTTGAGAAGGAACNTNAAAGAGGGGANGGGGAACA		900
QY	901	ATTTGTGAAGAAGAAATNCGGNGGGAATTTTGGCCCTTCAAGGGGAANAATGGCCCTGG		960
Db	901	ATTTGTGAAGAAGAAATNCGGNGGGAATTTTGGCCCTTCAAGGGGAANAATGGCCCTGG		960
QY	961	GGGGTAAAAAGANGGCCACCCCTGGGGGTGGTGGGAT	994	
Db	961	GGGGTAAAAAGANGGCCACCCCTGGGGGTGGTGGGAT	994	

RESULT 2
US-09-960-352-5678
; Sequence 5678, Application US/09960352

```

; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298) C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5678
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 24-LIB34-074-Q1-E1-F7
US-09-960-352-5678

```

Query Match	32.5%	Score 322.8;	DB 10;	Length 432;
Best Local Similarity	88.7%;	Pred. No. 4e-75;		
Matches 360; Conservative	0;	Mismatches 44;	Indels 2;	Gaps 1;

[illegible]

OY		597	GGACACCCTAGGCGCCTACAGTCTTGCTCAGTGTGCAGAACTCTGCTGGCAGCTGAGA	656
Db		242	GGGCACCCACTTGAGCTACAGTCTGGCTCAGTGC GTTGA ACTCTGCTGGCATCTGAGA	301
OY		657	GGGGAAGCCGGAAAAAGAGCAAAGTTCTGTGTCAAAAGTGCGCTCTGCNGCATTAATTAN	716
Db		302	GGGGAAGCCGGAAA - AAGCAAAGTTCTGTGTCAAAAGTTGCTCTGCAGCATATATAG	359
OY		717	GCATTGGAGGA CTGTG G TTG TAA CACTCTACAAGATGGGGTTTTTC	762
Db		360	GCATCGGAGAGCTGTCTGTTTACGCTGTACACAAGATGGGCTTTCC	405

RESULT 3

```

US-09-960-352-7555
? Sequence 7555, Application US/09960352
? Patent No. US20020137139A1
? GENERAL INFORMATION:
? APPLICANT: Warren, Wesley C.
? APPLICANT: Tao, Nengbing
? APPLICANT: Byatt, John C.
? APPLICANT: Machialagan, Nagappan
? TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
? TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
? FILE REFERENCE: 16511.006/37-21(10298)C
? CURRENT APPLICATION NUMBER: US/09/960,352
? CURRENT FILING DATE: 2001-09-24
? NUMBER OF SEQ ID NOS: 15112
? SEQ ID NO 7555
? LENGTH: 346
? TYPE: DNA
? ORGANISM: Bos taurus
? OTHER INFORMATION: Clone ID: 32-LIB34-076-Q1-E1-H7
US-09-960-352-7555

```

Query Match	25.5%	Score 253;	DB 10;	Length 346;
Best Local Similarity	86.8%	Pred. No. 1e-56;		
Matches 289; Conservative	0;	Mismatches 42;	Indels 2;	Gaps 1;

OY		430	ATATTGACGTAATAGAACTTCACGATTGCCTTTTCTACCAACGAACCTCCTTACTTATGAAG	489
Db		14	ATCCGGACATAATAGAACTTCACGATTGCCTTTTCTGCCAATGAACCTTATTACTTAGCGAAG	73
OY		490	CAC TGG G A C T C T G T C C A G A A G G A C A A G G T G C A A C G C T G G T T G A T A G A G G A T T A T A C A T	549
Db		74	C A C T G G G A C T G T G T C C A G A A G G T C A A G G T G A A A A C T G G T T G A A G A G G A G A T T A T A C T T	133
OY		550	A T G G A G G A A A G T G G G T C A T A A T C T A G T G T G G A C T G A T T T C A A A G G G A C A C C C A C T A G	609
Db		134	A T G G A G G A A A G T G G G T C A T A A T C C T A G T G G C G G A T T A T T T C A A A G G G G C A C C C A C T T G	193
OY		610	G C G C T A C A G G T C T T G C T C A G T G T G C A A A C T C T G C T G G C A G T G A G A G G G A A C C G G A A	669
Db		194	G A C C T A C A G G T C T G G C T C A G T G C G T T G A A C T C T G C T G G C A T C T G A G A G G G G A A C C C G G A A	253
OY		670	A A G A G G C A A A G T T C C T G T G T G C A A A G G T G G C T C T G C N G C A T A T T A N G C A T T G A G A G A A C	729
Db		254	A - - A A G G C A A G T T C C T G T G T G C A A A G G T T G C T C T G C A G C A T A T A T A T A G G C A T G S A T G A A C	311
OY		730	T G T G G T T G T A C A C A C T C T A C A A G A T G G G G T T T T C	762
Db		312	T G T G G T T G T A C G C T G T A C A A G A T G G G C T T T C C	344

RESULT 4

US-09-960-352-647
 / Sequence 647, Application US/09960352
 / Patent No. US20020137139A1
 / GENERAL INFORMATION:
 / APPLICANT: Warren, Wesley C.
 / APPLICANT: Tao, Nengbing
 / APPLICANT: Byatt, John C.
 / APPLICANT: Mathialagan, Nagappan

```

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 647
; LENGTH: 314
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 03-LIB34-050-Q1-E1-A11
US-09-960-352-647

```

Query Match

Query Match	23.9%	Score 238	DB 10	Length 314
Best Local Similarity	86.9%	Pred. No. 9.1e-53		
Matches 273	Conservative 0	Mismatches 40	Indels 1	Gaps 1
QY	358	AAATGGTTGGCTTGATATGAGTAAGAAGCTGCAGAAATGCTATGAGAAATCTGGCC	417	
Db	1	AAATGGTTGGTTTGATATGAGTAAGAAGCTGCCAGAAAGTGCTATGAGAAATCTGGCC	60	
QY	418	TGACACCAAAATGATATTGACGTAATAGAACTTCACGATTGCTTTCTACCAACGAACCTCC	477	
Db	61	TGAGACCAAGTGATATTGACGTAATAGAACTTCACGAACGCTTGCGCCAAATGAACCTTA	120	
QY	478	TTACTTATGAAGCACCTGGGACTCTGTCCAGAAGGACAAAGGTGCAACGCTGGTGTATAGAG	537	
Db	121	TTACTTACGAAGCACTGGGACTGTGTCCAGAATGTCAAGGTGGAATACTGGTGTGAAGAAT	180	
QY	538	GAGATAATACATATGAGAGAAAGTGGGTCATAAATCCTAGTGGTGAAGTTCATTTCAAAGG	597	
Db	181	GAGATAATACTTATGAGAGAAAGTGGGTCATAAATCCTAGTGGCGGATTAATTTCAAAGG	240	
QY	598	GACACCCCACTAGCGCGTACAGGCTTGTCTCACTGTGCAGAACTCTGCTGGCAGC-TGAGA	656	
Db	241	GGCACCCCACTTGAGCTACAGGCTTGCGCTCACTGTGAACTCTGCTGGCATCTTGAGA	300	
QY	657	GGGGAAGCCGGAAA 670		
Db	301	CGGGGAAGCCGGAAA 314		

RESULT 5

US-10-156-761-5843
; Sequence 5843, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

APPLI

```

1  APPLICANT:  HORIKAWA, HIROSHI
2  APPLICANT:  SHIBA, TADAYOSHI
3  APPLICANT:  SAKAKI, YOSHIYUKI
4  APPLICANT:  HATTORI, MASAHIRA
5  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
6  FILE REFERENCE: 249-262
7  CURRENT APPLICATION NUMBER: US/10/156,761
8  CURRENT FILING DATE: 2002-05-29
9  PRIOR APPLICATION NUMBER: JP 2001-204089
10 PRIOR FILING DATE: 2001-05-30
11 PRIOR APPLICATION NUMBER: JP 2001-272697
12 PRIOR FILING DATE: 2001-08-02
13 NUMBER OF SEQ ID NOS: 15109
14 SEQ ID NO 5843
15 LENGTH: 1236
16 TYPE: DNA
17 ORGANISM: Streptomyces avermitilis
18 FEATURE:
19 NAME/KEY: CDS
20 LOCATION: (1)..(1236)
21 US-10-156-761-5843

```

Query Match

19.78; Score 196.2; DB 14; Length 1236;

Best Local Similarity 55.1%; Pred. No. 2.2e-41;
Matches 402; Conservative 0; Mismatches 325; Indels 2; Gaps 1;

```
QY 16 CTCAGATGTTGGGTATGCTGGAAGAACATATGAAAAATATGAAACAAAATTGAAC 75
   |||||
Db 476 CGCAGATCTTCGGCGACGCGCGCGCGAGACATGAGAAAGTACGGGACGACCGAGCGC 535

QY 76 ACTTGCAGAAAATTGGATGAAAAATCATAAACATTCACTTAATTAACCCGTATCCAGT 135
   |||||
Db 536 AACTGCGCGCGGTGCGGGCCAGAACACCGGCACTCGCGCACAAACCCGTACGCCAGT 595

QY 136 TCCAGATGATACAGTTAGATGAAGTATGCGATCTAAAGAGTTTGTATTTTGA 195
   |||||
Db 596 TCCAGAGCGGTACACCGTCGACGAGATCCTCGCGTCCGCCCGGTGACACCGCCGCTGA 655

QY 196 CTATCTTACAATGTTGTCCCACTTCAGATGCTGCTGCAGCAGCAATTTGGCCAGTGAAG 255
   |||||
Db 656 CCAAGCTTCAAGTCTCACCACCTCGGACGAGCGCGCGCGCGCGGTGCTCTCCGAAC 715

QY 256 CATTGTACAGAGATATGCGCTGCAATCCAAAGCTGTGGAATTTTGGCACAAGAAATGA 315
   |||||
Db 716 GCTTCGCGACGCGGACGCGGCTGGGGACAGGCTGTGAGATGTCGCGCAGCGGATGA 775

QY 316 TGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAATGCTTGCTTGATA 375
   |||||
Db 776 CGACGGAACAGGAGAGTCTTCGCGTGGGGTCTTCGATCGACGTCGTCGCGCGCGCA 835

QY 376 TGAAGTAAGAAGCTGCAAGAAAATGCTATGAGAAATCTGGCTGACCAACATGATATG 435
   |||||
Db 836 TGTCCGCGGAGGCGCGCGCGCGGTGTACGAGCGTGCAGGCTCGGATCGAGAGCTCG 895

QY 436 ACCTAATAGAACTTCACGATGCTTTTCTACCAACGAACTCCTTACTTATGAAGCACTGG 495
   |||||
Db 896 ACGTGTGGAACCTGCAACGACTGCTTCTGATCAACGAGCTGCTGACGTAACGAGCACTCG 955

QY 496 GACTCTGTCCAGAGGACAGCAAGGTGCAACGCTGCTGATAGAGAGATAATACATATGAG 555
   |||||
Db 956 GCATGTGCTGCGCGCGGAGTCCGGGAAGCTGTGAGTCGGGAGCGACGACTACGCGCG 1015

QY 556 GAAAGTGGCTCATAAATCCTAGTGTGGAAGTTCGAAAGGACACCCACTAGGCGCTA 615
   |||||
Db 1016 GCCGTGGGTCTGAACCCGTCGCGCGGCTGATCTCCAAGGGGATCCGCTGGCGCCA 1075

QY 616 CAGCTCTGCTCAGTGTGAGAACTGCTGCGAGCTGAGAGGGGAAACCGAAAAAGAG 675
   |||||
Db 1076 CGGATATGCCAGATCGCCGAGCTGACCTGCGCGGCGAGGCGGG--GGCAGC 1133

QY 676 CAAAGTCTCTGTCGCAAGAGTGGCTCTGCGCATATATTANGCATTTGAGAGACTGTGT 735
   |||||
Db 1134 CCAGGTCCCGGGCGCGCGGGTGGGTCTCGCGCACACATCGGACTGGCGGAGCGCGGT 1193

QY 736 TGTAACT 744
   |||||
Db 1194 GGTGACGCT 1202
```

RESULT 6

US-10-156-761-1

; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 19.7%; Score 196.2; DB 14; Length 9025608;
Best Local Similarity 55.1%; Pred. No. 3e-39;
Matches 402; Conservative 0; Mismatches 325; Indels 2; Gaps 1;

```
QY 16 CTCAGATGTTGGGTATGCTGGAAGAACATATGAAAAATATGAAACAAAATTGAAC 75
   |||||
Db 7075437 CGCAGATCTTCGGCGACGCGCGCGCGAGCACATGAGAAAGTACGGGACGACCGAGCGC 7075496

QY 76 ACTTGCAGAAAATTGGATGAAAAATCATAAACATTCACTTAATTAACCCGTATCCAGT 135
   |||||
Db 7075497 AACTGCGCGCGGTGCGGGCCAGAACACCGGCACTCGCGCACAAACCCGTACGCCAGT 7075556

QY 136 TCCAGATGAATACAGTTTACATGAAGTATGCGATCTTAAAGAGTTTGTATTTTGA 195
   |||||
Db 7075557 TCCAGAGCGGTACACCGTCGACGAGATCCTCGCGTCCGCCCGGTGACACCGCCGCTGA 7075616

QY 196 CTATCTTACAATGTTGTCCCACTTCAGATGCTGCTGCAGCAGCAATTTGGCCAGTGAAG 255
   |||||
Db 7075617 CCAAGCTCCAGTGTCTACCCCACTTCGAGCAGAGCGCGCGCGCGTGTCTCTCCGAA 7075676

QY 256 CATTGTACAGAGATATGCGCTGCAATCCAAAGCTGTGGAATTTTGGCACAAGAAATGA 315
   |||||
Db 7075677 GCTTCGCGACGCGGACGCGGCTGGGGACAGAGTGTGAGATCTCGCGACGCGATGA 7075736

QY 316 TGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTATTAATGTTGCTTGATTA 375
   |||||
Db 7075737 CGACGGAACAGGAGAGTCTTCGCGTGGGGTCTTCGATCGACGTCGCGCGCGCA 7075796

QY 376 TGAAGTAAGAAGCTGCAAGAAAATGCTATGAGAAATCTGCGCTGACACCAATGATATTG 435
   |||||
Db 7075797 TGTGCGGGAGGCGCGCGCGGGTGTACGAGCGTGCAGGCTCGGATCGAGGAGCTCG 7075856

QY 436 ACCTAATAGAACTTCACGATGCTTTTCTTACCAACGAACTCCTTACTTATGAAGCACTGG 495
   |||||
Db 7075857 ACGTGTGGAACCTGCAACGACTGCTTCTGATCAACGAGCTGCTGACGTAACGAGCACTCG 7075916

QY 496 GACTCTGTCCAGAGGACAGCAAGGTGCAACGCTGCTGATAGAGAGATAATACATATGAG 555
   |||||
Db 7075917 GCATGTGCTGCGCGGGAGTCCGGAAAGCTGTGAGTCGGGAGCGACGACTACGCGC 7075976

QY 556 GAAAGTGGGTATTAATCCTAGTGTGAGTGTGATTTCAAGGAGACCCACTAGGCGCTA 615
   |||||
Db 7075977 GCCGTGGGTCTGAACCCGTCGCGCGGGCTGATCTCCAAGGGGATCCGCTGGCGCCA 7076036

QY 616 CAGCTCTGCTCAGTGTGAGAACTGCTGCTGCGAGCTGAGAGGGGAAACCGGAAAGAG 675
   |||||
Db 7076037 CGGATATCGCCAGATCGCCGAGCTGACCTGCGAGCTGCGCGGAGGCGGG--GGCAGC 7076094

QY 676 CAAAGTCTCTGTCGCAAGAGTGGCTCTGCGCATATATTANGCATTTGAGAGAACTGTGT 735
   |||||
Db 7076095 CCAAGTCCCGGGCGCGCGGGTGGGTCTCGCGCACACATCGGACTGGCGGAGCGCGGT 7076154

QY 736 TGTAACT 744
   |||||
Db 7076155 GGTGACGCT 7076163
```

RESULT 7

US-09-960-352-2946

```
/ Sequence 2946, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Machlalan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 2946
/ LENGTH: 412
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 13-LIB34-005-Q1-E1-D1
US-09-960-352-2946
```

```
Query Match      19.1%; Score 189.8; DB 10; Length 412;
Best Local Similarity 86.0%; Pred. No. 5.9e-40;
Matches 221; Conservative 0; Mismatches 34; Indels 2; Gaps 1;
```

```
QY 506 AGAAGACAAGGTGCAACCGTGGTTGATAGAGAGATATACATATGAGGAAAGTGGT 565
DB 116 AGCAGGTCAAGGTGAAAACTGGTTGAAAGAGAGATATATCTTATGAGGAAAGTGGT 175
QY 566 CATTAATCCTAGTGTGACTGATTCAAGGAGACACCACTAGGCGCTACAGGTTCTGC 625
DB 176 CATTAATCCTAGTGTGCGGATTAATTTCAAGGGGACCCCACTTGAGCTACAGGTTGCG 235
QY 626 TCAGTGTGCAAGACTCTGCTGCGAGCTGAGAGGGAAGCCGAAAGAGCAAAAGTTCT 685
DB 236 TCAGTGTGCTGAACTCTGCTGCGATCTGAGAGGGAAGCCGAAAG--AAGCAAGTTCT 293
QY 686 GGTGCAAGGTGCTCTGCGCATTAATTANGCATTTGAGAGAACTGTGTTGTAACACTC 745
DB 294 GGTGCAAGGTGCTCTGCGCATTAATATAGCGCTCGAGAGACTGTGTTGTAACACTC 353
QY 746 TACAAGATGGGTTTC 762
DB 354 TACAAGATGGGTTTC 370
```

RESULT 8

```
US-09-878-574-8890
/ Sequence 8890, Application US/09878574
/ Patent No. US20020110548A1
/ GENERAL INFORMATION:
/ APPLICANT: Byrum, Joseph R.
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Thompson, Michael D.
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ FILE REFERENCE: 38-21(15401)B
/ CURRENT APPLICATION NUMBER: US/09/878,574
/ CURRENT FILING DATE: 2001-12-21
/ PRIOR APPLICATION NUMBER: 09/333,535
/ PRIOR FILING DATE: 1999-06-14
/ NUMBER OF SEQ ID NOS: 15775
/ SEQ ID NO 8890
/ LENGTH: 249
/ TYPE: DNA
/ ORGANISM: Glycine max
/ OTHER INFORMATION: Clone ID: 701101811H1
US-09-878-574-8890
```

```
Query Match      9.3%; Score 92; DB 10; Length 249;
Best Local Similarity 63.1%; Pred. No. 3.4e-14;
Matches 159; Conservative 0; Mismatches 90; Indels 3; Gaps 1;
```

QY 235 CAGCAATTTTGGCCAGTGAAGCAATTGTACAGAAGTATGGCTGCAATCCAAAGCTGTG 294

```
DB 1 CTGCTATCTGTGTCAGTGCAGAGAGTTGTGTGAAACCAATTTGCCATCACAAGCTGTG 60
QY 295 AAATTTGGCACAAGAAATGATGACTGATTTGGCCAGCTGTTGAAGAAAAAGCATTA 354
DB 61 ACATTCCTGCTGTGACATGCGCCACAGACACTGAACCTTCTTCAAGGAC--AGCTGTA 117
QY 355 TTAATGTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTG 414
DB 118 TCAAGATGGCTGTTTGACATGCGCAATTTGGCTGCAAGTAAAGCTTACAAGATGCTG 177
QY 415 GCCTGACACCAATGATTTGACGTAATAGAACTTCAGATTGCTTTCTACCAAGAAC 474
DB 178 GCGTCACCGCTAACGATGTTCAAGTATTTGAACCTTCAGATTGTTTCAAGTCAACGAAC 237
QY 475 TCCTTACTTATG 486
DB 238 TCGTTACATACG 249
```

RESULT 9

```
US-10-102-524-364
/ Sequence 364, Application US/10102524
/ Publication No. US20030109434A1
/ GENERAL INFORMATION:
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Gordon, Brian
/ APPLICANT: Harlocker, Susan L.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ FILE REFERENCE: 210121.572
/ CURRENT APPLICATION NUMBER: US/10/102,524
/ CURRENT FILING DATE: 2002-03-19
/ NUMBER OF SEQ ID NOS: 1863
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 364
/ LENGTH: 553
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 539
/ OTHER INFORMATION: n = A,T,C or G
US-10-102-524-364
```

```
Query Match      9.0%; Score 89.4; DB 14; Length 553;
Best Local Similarity 71.9%; Pred. No. 2.5e-13;
Matches 205; Conservative 0; Mismatches 62; Indels 18; Gaps 7;
```

```
QY 701 CTGCGCATTAATTANGCATTTGAGGAGAACTGTGTTGTAACACTCTACAAGATGGGTTT 760
DB 1 CTGACGATTAATTAGCATTTGAGAGAGCTGTGTTGTAACACTCTACAAGATGG--TT 58
QY 761 TCCCGGAAGCCGCACTCTTTAGAACTCATCAAAATGAAGCCNGTTCCAACCAAGC 820
DB 59 TTCCGGAAGCCGCACTT-CTTTAGAACTCATCAAAATGA--GCTGTTCCAACCAAGC 114
QY 821 TCTGCAAGTATNGTTTAAAGNAAATCTNGTTTAAAGNGGATTTGAGAGAAACNT 880
DB 115 TCTGCAAGTATGATTTAAGCAAAATCTGTTTAA--GGAATTGAGAGAAACTTG 172
QY 881 NAAAGAGGAAGGGGAACAATTTGTGAAGAAAAATNGGNGGAATTTTGCCCTTCA 940
DB 173 A---AGAGGAAGGGGAACAGTTTGTGAAGAAAAATCGGTGT-----ATTTTGCTTCA 223
QY 941 AGGGGAANAATGGCCCTGGGGGGTAAAGANGGCCACCTGGGG 985
DB 224 AGGTGAAG-ATGGCCCTGGGGGGTAAAGAGGCCACCTGGGTGTG 267
```

RESULT 10

```
US-10-027-632-34138/C
; Sequence 34138, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34138
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(938)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-34138

Query Match          6.8%; Score 67.2; DB 12; Length 938;
Best Local Similarity 83.3%; Pred. No. 2.5e-07;
Matches 75; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY      209 TTGTCCCACTTCAGATGTGCTGCAGCAGCAATTTGGCCAGTGAAGCATTGTACAGAA 268
      |||||||
Db      652 TAGTCCTACCTCAGATGTGCTGCAGTGCACAAATTTGGCCAGKGAAGCATTGTGACAGGA 593

QY      269 GTATGGCCTGCAATCCAAAGCTGTGGAAT 298
      |||||||
Db      592 GTATGGCCTGCAACCCAAAGAGTTCACAT 563

RESULT 11
US-10-027-632-34139/C
; Sequence 34139, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34139
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(938)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-34139

Query Match          6.8%; Score 67.2; DB 12; Length 938;
Best Local Similarity 83.3%; Pred. No. 2.5e-07;
Matches 75; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY      209 TTGTCCCACTTCAGATGTGCTGCAGCAGCAATTTGGCCAGTGAAGCATTGTACAGAA 268
      |||||||
Db      652 TAGTCCTACCTCAGATGTGCTGCAGTGCACAAATTTGGCCAGKGAAGCATTGTGACAGGA 593

QY      269 GTATGGCCTGCAATCCAAAGCTGTGGAAT 298
      |||||||
Db      592 GTATGGCCTGCAACCCAAAGAGTTCACAT 563

RESULT 12
US-10-027-632-34138/C
; Sequence 34138, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34138
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(938)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-34138

Query Match          6.8%; Score 67.2; DB 13; Length 938;
Best Local Similarity 83.3%; Pred. No. 2.5e-07;
Matches 75; Conservative 1; Mismatches 14; Indels 0; Gaps 0;
```

```
RESULT 13
US-10-027-632-34139/c
; Sequence 34139, Application US/10027632
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   PRIOR FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
;   PRIOR FILING DATE: 1999-09-28
;   PRIOR APPLICATION NUMBER: US 60/146,002
;   PRIOR FILING DATE: 1999-08-09
;   NUMBER OF SEQ ID NOS: 325720
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 34139
;   LENGTH: 938
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(938)
;   OTHER INFORMATION: n = A,T,C or G
US-10-027-632-34139

Query Match      6.8%; Score 67.2; DB 13; Length 938;
Best Local Similarity 83.3%; Pred. No. 2.5e-07;
Matches 75; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

QY      209 TTGTCCCACTTCAGATGCTGCTGCAGCAGCAATTTGGCCAGTGAAGCATTTGTACAGAA 268
      |||||
Db      652 TAGTCCTACCTCAGATGGTGTCTGCAGTGGCAATTTGGCCAGKGAAGCATTTGCACAGCA 593

QY      269 GTATGGCCTGCAATCCAAAGCTGTGGAAT 298
      |||||
Db      592 GTATGGCCTGCAACCCAAGAGTTCCACAT 563

RESULT 14
US-10-027-632-103694/c
; Sequence 103694, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   CURRENT FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103694
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-103694

Query Match      4.3%; Score 42.6; DB 12; Length 2432;
Best Local Similarity 51.9%; Pred. No. 1.4;
Matches 96; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY      305 ACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAGAGAAAAAGCATTTAATAATGCT 364
      |||||
Db      416 ACATGAAATATCTAGCTTAAATAGATAATTGATAAAAAAGGTTTCAGATATGCT 357

QY      365 TGGCTTTGATATGAGTAAGAAGCTGCAGAGAAATGCTATGAGAAATCTGGCCTGACACC 424
      |||||
Db      356 TTTCATGTAGTATCACAAGCATCTGCCAGAGTAAGCAATTAGTATCTCTTTGAAAAAT 297

QY      425 AATGATATTGACGTAATAGAACTTCAGGATTCCTTTCTACCAAGCAACTCCTTACTTA 484
      |||||
Db      296 TAAAAAAGAAACAGAAAAATCCATTATCTTTTTTTATTTTCCAAACTGCTTCAGAA 237

QY      485 TGAAG 489
      |||||
Db      236 AGAAG 232
```

```
RESULT 15
US-10-027-632-103695/c
; Sequence 103695, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;   APPLICANT: Wang, David G.
;   TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   FILE REFERENCE: 108827.129
;   CURRENT APPLICATION NUMBER: US/10/027,632
;   CURRENT FILING DATE: 2002-04-30
;   PRIOR APPLICATION NUMBER: US 60/218,006
;   PRIOR FILING DATE: 2000-07-12
;   PRIOR APPLICATION NUMBER: US 60/198,676
;   PRIOR FILING DATE: 2000-04-20
;   PRIOR APPLICATION NUMBER: US 60/193,483
;   PRIOR FILING DATE: 2000-03-29
;   PRIOR APPLICATION NUMBER: US 60/185,218
;   PRIOR FILING DATE: 2000-02-24
;   PRIOR APPLICATION NUMBER: US 60/167,363
;   PRIOR FILING DATE: 1999-11-23
;   PRIOR APPLICATION NUMBER: US 60/156,358
;   PRIOR FILING DATE: 1999-09-28
;   PRIOR APPLICATION NUMBER: US 60/146,002
;   PRIOR FILING DATE: 1999-08-09
;   NUMBER OF SEQ ID NOS: 325720
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 103695
;   LENGTH: 2432
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103695

Query Match      4.3%; Score 42.6; DB 12; Length 2432;
Best Local Similarity 51.9%; Pred. No. 1.4;
Matches 96; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY      305 ACAAGAAATGATGACTGATTTGCCAAGCTCGTTGAGAGAAAAAGCATTTAATAATGCT 364
      |||||
Db      416 ACATGAAATATCTAGCTTAAATAGATAATTGATAAAAAAGGTTTCAGATATGCT 357
```

QY	365	TGGCTTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGACACC	424
Db	356	TTTCATGTAAAGTATCACAAAGCATCTGCCAGAGTAGCAATTAGTATCTCTTTGAAAT	297
QY	425	AAATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTACCAACGAACCTTACTTA	484
Db	296	TAAAAAAGGAAACAGAAAAATCCATTATCTTTTATTATTTCCTCAAACTGCTTCAGAA	237
QY	485	TGAAG	489
Db	236	AGAAG	232

Search completed: November 27, 2003, 12:53:07
Job time : 384.822 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 06:00:22 / Search time 73.3681 Seconds
(without alignments)
5979.910 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 994

Sequence: 1 CTCACCCAGTGTCTCTCCAG.....CCACCTGGGTGTGGTGGAT 994

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents, NA: *
2: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	978	98.4	994	3	US-08-896-164-21
2	214	21.5	4411529	3	US-09-103-840A-1
3	199.6	20.1	4403765	3	US-09-103-840A-2
4	92	9.3	1664976	4	US-08-916-421B-1
5	49.2	4.9	7218	1	US-08-232-463-14
6	39.8	4.0	1644	1	US-08-785-048-4
7	39.8	4.0	1644	2	US-08-996-799-4
8	39.8	4.0	1662	1	US-08-785-048-1
9	39.8	4.0	1662	2	US-08-996-799-1
10	39	3.9	580073	4	US-08-545-528D-1
11	38.2	3.8	1368	4	US-09-328-352-1034
12	37.8	3.8	1664976	4	US-08-916-421B-1
13	37.4	3.8	1353	4	US-09-252-991A-13040
14	37.4	3.8	1410	4	US-09-252-991A-12438
15	35.8	3.6	19124	2	US-08-487-826B-13
16	35.8	3.6	1830121	4	US-09-557-884-1
17	35.8	3.6	1830121	4	US-09-643-990A-1
18	35	3.5	1464	3	US-08-605-150A-3
19	34.8	3.5	393	4	US-09-328-352-826
20	34.8	3.5	2079	3	US-09-381-849-4
21	34.8	3.5	2193	4	US-09-328-352-316
22	34.6	3.5	1301	1	US-08-431-080-1
23	34.6	3.5	1301	1	US-08-431-080-4
24	34.6	3.5	1301	2	US-08-938-534-1
25	34.6	3.5	1301	2	US-08-938-534-4
26	34.6	3.5	1301	4	US-09-345-294-1
27	34.6	3.5	1301	4	US-09-345-294-4

28	34.6	3.5	1490	3	US-08-605-150A-5	Sequence 5, Appli
29	34.4	3.5	580073	4	US-08-545-528D-1	Sequence 1, Appli
30	34.2	3.4	599	3	US-09-328-111-219	Sequence 219, App
31	33.8	3.4	572	3	US-09-328-111-87	Sequence 87, Appl
32	33.8	3.4	614	3	US-09-328-111-111	Sequence 111, App
33	33.8	3.4	10411	4	US-08-961-527-89	Sequence 89, Appl
34	33.6	3.4	22067	4	US-09-820-001-3	Sequence 3, Appli
35	33.6	3.4	37948	3	US-09-251-645-11	Sequence 11, Appl
36	33.4	3.4	3663	3	US-09-499-884-11	Sequence 11, Appl
37	33.4	3.4	15016	4	US-09-601-198-60	Sequence 60, Appl
38	33.2	3.3	582	4	US-09-107-532A-1201	Sequence 1201, Ap
39	33.2	3.3	3627	1	US-08-447-500-23	Sequence 23, Appl
40	33.2	3.3	3627	1	US-08-454-097-23	Sequence 23, Appl
41	33.2	3.3	3627	1	US-08-453-866-23	Sequence 23, Appl
42	33.2	3.3	3627	3	US-08-185-359-23	Sequence 23, Appl
43	33.2	3.3	3628	1	US-08-468-036-17	Sequence 17, Appl
44	33.2	3.3	3628	2	US-08-376-843-17	Sequence 17, Appl
45	33.2	3.3	9873	4	US-09-328-352-1360	Sequence 1360, Ap

ALIGNMENTS

RESULT 1
US-08-896-164-21
; Sequence 21, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 994 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-896-164-21

Query Match 98.4%; Score 978; DB 3; Length 994;
Best Local Similarity 100.0%; Pred. No. 2.2e-271;
Matches 994; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCACCCAGTGTCTCTCAGATGTTGGTATGCTGAAAGAACATATGAAATAATG 60
Db 1 CTCACCCAGTGTCTCTCAGATGTTGGTATGCTGAAAGAACATATGAAATAATG 60
QY 61 GAACAAATTTGAACACTTTCGAAATAATTTGATGGAATAATCATTAACATTCACTTAATA 120

```
|||||
Db      61 GAACAAAATTTGAACCTTTGCAAAAATTTGATGAAAAATCATTAACATTGAGTTAATA 120
QY      121 ACCCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGATGCACTTAAGAAG 180
Db      121 ACCCGTATCCAGTTCACAGATGAATACAGTTAGATGAAGTGATGCACTTAAGAAG 180
QY      181 TTTTGAATTTTGAATCTTACAAATGTTGCCACTTCAGATGCTGACAGACAA 240
Db      181 TTTTGAATTTTGAATCTTACAAATGTTGCCACTTCAGATGCTGACAGACAA 240
QY      241 TTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCTGCAATCCAAGCTGTGAATTT 300
Db      241 TTTGGCCAGTGAAGCATTTGTACAGAGTATGGCCTGCAATCCAAGCTGTGAATTT 300
QY      301 TGGCACAAGAAATGATGACTGATTGCAAGCTCGTTGAAGAAAAAGCATTATTA 360
Db      301 TGGCACAAGAAATGATGACTGATTGCAAGCTCGTTGAAGAAAAAGCATTATTA 360
QY      361 TGGTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA 420
Db      361 TGGTTGGCTTTGATATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGA 420
QY      421 CACCAAAATGATATGACGTAATAGAACTTCAAGATTGCTTTTCTAACGAACCTCTTA 480
Db      421 CACCAAAATGATATGACGTAATAGAACTTCAAGATTGCTTTTCTAACGAACCTCTTA 480
QY      481 CTTATGAAGCACTGGGACTCTGTCCAGAAGACAAGGTGCAACGCTGTGATAGAGAG 540
Db      481 CTTATGAAGCACTGGGACTCTGTCCAGAAGACAAGGTGCAACGCTGTGATAGAGAG 540
QY      541 ATAATACATATGAGAGAAAGTGGGTCTAATAATCCTAGTGTGACTGATTTCAAAGGAC 600
Db      541 ATAATACATATGAGAGAAAGTGGGTCTAATAATCCTAGTGTGACTGATTTCAAAGGAC 600
QY      601 ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAAGAACTCTGCTGGCAGCTGAGAGGG 660
Db      601 ACCCACTAGGCGCTACAGGCTTGTCTCAGTGTGCAAGAACTCTGCTGGCAGCTGAGAGGG 660
QY      661 AAGCCGGAAGAGGCAAAAGTTCCTGTGTCAAAAGTGGCTCTGCGCATATTTANGCAT 720
Db      661 AAGCCGGAAGAGGCAAAAGTTCCTGTGTCAAAAGTGGCTCTGCGCATATTTANGCAT 720
QY      721 TGGAGAACTGTGTTGTAAACACTCTACAGATGGGGTTTCCGGAAGCCGACAGTTCC 780
Db      721 TGGAGAACTGTGTTGTAAACACTCTACAGATGGGGTTTCCGGAAGCCGACAGTTCC 780
QY      781 TTTTAGACTCATCAAAATTTGAAGCCTTCCAAACCAAGCTCTGCAAGTATNGTTAA 840
Db      781 TTTTAGACTCATCAAAATTTGAAGCCTTCCAAACCAAGCTCTGCAAGTATNGTTAA 840
QY      841 NGNAAATCTNGTTTAAAGNGATTTGAGAAGGAAACNTNAAAGGGAANGGGAACA 900
Db      841 NGNAAATCTNGTTTAAAGNGATTTGAGAAGGAAACNTNAAAGGGAANGGGAACA 900
QY      901 ATTGTGAAAGAAAAATNCGNGGAAATTTTGCCCTTCAAGGGGAANAATGGCCCTGG 960
Db      901 ATTGTGAAAGAAAAATNCGNGGAAATTTTGCCCTTCAAGGGGAANAATGGCCCTGG 960
QY      961 GGGGTAAAGANGGCCACCTGGGGTGTGGGAT 994
Db      961 GGGGTAAAGANGGCCACCTGGGGTGTGGGAT 994
```

```
RESULT 2
US-09-103-840A-1/c
; Sequence 1, Application us/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
```

```
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
```

```
Query Match      21.5%; Score 214; DB 3; Length 4411529;
Best Local Similarity 57.1%; Pred. No. 3.7e-50;
Matches 428; Conservative 0; Mismatches 317; Indels 5; Gaps 2;
```

```
QY      6 CCAATTGCTCCTCAGATGTTGGGTATGCTGGAAGAAGACATATGAAAAATATGAACA 65
Db      3099716 CCGGTGGCCGCCGTGATGTTGGGGGGCGCCGCCGGAACATGAAGAATACGGCAC 3099657
QY      66 AAAATTGAACACTTTGCCAAAATTGGATGAAAAATCATAAACATTCAATTAAACCCG 125
Db      3099656 ACCGGGAGCATTTTCGCGAAGATCGGCTACAGAAGACCAAGCACTCGGTCAACAACCCG 3099597
QY      126 TATTCCTCAGTTCGAAGATGATACAGTTAGATGAAGTATGCGATCTAAGAAGTTT 185
Db      3099596 TATGACAGTTTCAGAGCAGATACACCTTGACGACATCTTGCCCTCAAGATGATTTCC 3099537
QY      186 GATTTTGTACTATCTTACATGTTGTGCCACTTCAGATGCTGACAGCAATTTTG 245
Db      3099536 GACCCGCTGACCAAAATGCAAGTGTCTCCCACTCCGAGGGGTGCGCGGTGTGCTG 3099477
QY      246 GCCAGTGAAGCATTTGTACAGAAGTATGCGCTGCAATCCAAGCTGTGAAATTTGGCA 305
Db      3099476 GCCAGTGAAGATTACCTGGCCAACAACACTTGCCGGCGGGCTGTGAAATCGTGGG 3099417
QY      306 CAGAAATGATGACTGATTTGCCAAGCTCGTTTGAAGAAAAAGCATTTAATAATGTT 365
Db      3099416 CAGGCGATGACCAACCGACTTCCTTCCACCTTTGATG---GCAGTCCCCGAATATCATC 3099360
QY      366 GCGTTTGTATGAGTAAAGAGCTGCAAGAAATGCTATGAGAAATCTGGCCTGACACA 425
Db      3099359 GCGTACGACATGACTGTGCAAGCGGACACAACGGGTTTACAGCAATCCGACTCGGCCCG 3099300
QY      426 AATGATATTTGACGTAATAGAACTTCAAGATTGCTTTTCTACCAACGAACCTCTTACTAT 485
Db      3099299 AAGGACTTGGAGTATGAGCTGACGACTGCTTCTCAGCCAAAGAGTACTGCTTAC 3099240
QY      486 GAAGCACTGGACTCTGTCCAGAAGACCAAGGTGCAACGCTGTGATAGAGAGATAAT 545
Db      3099239 GAGCGCTCGGCTGTGGGGCCGGGAGAGCGCCGAGCTGATGACAGACAACGAGACC 3099180
QY      546 ACATATGAGAGAAAGTGGGTCTAATAATCCTAGTGTGAGCTGATTTCAAGGGACACCA 605
Db      3099179 ACCTACGGCGGACGCTGGGTCTCAACCATCCGGTGGCTGATCTCCAAGGGCATCCG 3099120
QY      606 CTAGCGCTTACAGTCTTCTCAGTGTGCAAGACTCTGCTGGCAGCTGAGAGGGGAAGCC 665
Db      3099119 CTGGGTGCGACGGGGTGGCGCAGTGGCGGAGCTGCACTGGCAGCTGCGCGGACCCG 3099060
QY      666 GGAAGAGAGCAAGTCTCTGTGCAAAAGGTGCTCTGCGCATTAATTANGCATTTGAG 725
Db      3099059 G--AGCGCGCAGGTGCAACAGTGACCGCGCGCTGCAACACACATCGGCTGGCG 3099002
QY      726 GAAGTGTGTTGTAACACTCTACAAGATGG 755
Db      3099001 GCGCGCGCGTGTACCGCATACCAACGGG 3098972
```

```
RESULT 3
US-09-103-840A-2/c
```

```
/ Sequence 2, Application US/09103840A
/ Patent No. 6294328
/ GENERAL INFORMATION:
/ APPLICANT: FLEISCHMAN, Robert D.
/ APPLICANT: WHITE, Owen R.
/ APPLICANT: FRASER, Claire M.
/ APPLICANT: VENTER, John C.
/ TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
/ TITLE OF INVENTION: TUBERCULOSIS
/ FILE REFERENCE: 24366-2007.00
/ CURRENT APPLICATION NUMBER: US/09/103,840A
/ CURRENT FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 4403765
/ TYPE: DNA
/ ORGANISM: Mycobacterium tuberculosis
/ FEATURE:
/ OTHER INFORMATION: CDC 1551
/ OTHER INFORMATION: "n" bases at various positions throughout the sequence
/ OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Query Match      20.1%; Score 199.6; DB 3; Length 4403765;
Best Local Similarity 56.8%; Pred. No. 5.1e-46;
Matches 426; Conservative 0; Mismatches 316; Indels 8; Gaps 3;
```

```
QY      6 CCAGTGTCTCCTCAGATGTTGGGTATGCTGGAAAAAGACATATGAAAAATATGAAACA 65
Db      3094392 CCGGTGGCGCGCGTGATGTTGGGGGGCGCGCGCGGAACACATGAAGAAATACGGCAC 3094333

QY      66 AAAATTGAACACTTTGCCAAAAATTGATGAAAAATCATTAACATTCAAGTTAATACCCG 125
Db      3094332 ACCGGGAGCATTTTCGGAAGATGGCTACAAAGAACCAAGCACTCGGTCAACAACCCG 3094273

QY      126 TATTTCCAGTTCAGATGAATACAGTTAGATGAAGTATGCACTTAAGAAGTTT 185
Db      3094272 TATGACACAGTTTCAGACGAATACACCTTGACGACATCTGGCCTCAAGATGATTTC 3094213

QY      186 GATTTTGTACTATCTTACAATGTTGCCACTTCAGATGTTGCTGCACAGCAATTTG 245
Db      3094212 GACCCGCTGACCAATTCAGTGTCTCCACCTCCGACGGGTGCGCGGTGCTG 3094153

QY      246 GCCAGTGAAGCATTTGTACAGAGATGCGCTGCAATCCAAAGCTGGAATTTGGCA 305
Db      3094152 GCCAGTGAAGGATTACCTGGCCCAACCAACCTTGCCGGGGGGGCTGTGAATCTCGGG 3094093

QY      306 CAAGAAATGATGACTGATTGTCAGCTCGTTGAAGAAAAAGCATTAATAATGTT 365
Db      3094092 CAGCGATGACCAACGACTTCGCTTCCACTTTGATG--GCAGTGCCCGCAATATCATC 3094036

QY      366 GCGTTGATATGAGTAAAGAGTGCAGAAAAATGCTATGAAGAAATCTGGCCTGACACA 425
Db      3094035 GCGTACGACATGACTGTGCAAGCGGCAACAAGGTTTACCAAGCATCCGACTCGGCCG 3093976

QY      426 AATGATATTGACGTAATAGAACTTCACGATTGCTTTTCTACCAAGCACTCCTTACTTAT 485
Db      3093975 AAGGACTTCGGAAGTATCGAGCTGCACGACTGCTTCTCAGCCCAAGACTACTGCTCTAC 3093916

QY      486 GAAGCACTGGGACTCTGTCCAGAAAGACAAGGTGCAACCGCTGTTGATAGAGGAGATAT 545
Db      3093915 GAGGCGCTCGGCTGTGCGGGCGGCGGAGGCGCCGAGCTGATCGACCAACCAAGACC 3093856

QY      546 ACATATGGAGGAAAGTGGGTCAATAATCTAGTGTGAGTGAATTTCAAAGGACACCCA 605
Db      3093855 ACCTACGGCGGACGCTGGGTCTCAACCATCCGGTGGGCTGATCTCCAAGGCCATCCG 3093796

QY      606 CTAGCGGCTACAGGCTCTGTCTCAGTGTGACAAGCTCTGCTGCGAGCTGAGAGGGGAAGCC 665
Db      3093795 CTGGGTGCGACGGGGTGGCGCAGTG--CGAGCTGACCTGGCAGCTGCGCGCACCGCC 3093739

QY      666 GAAAAAGAGGCAAAAGTTCTGTGTGCAAGAGTGGCTCTGCNGCATTAATTANGCATTTGAG 725
```

```
Db      3093738 G--AGCGCGCCAGGTCGACAAAGTGACCCGCCCGCTGTCAACACACATCGGGCTGGCG 3093681
QY      726 GAAGTGTGTTGTAACTCTTACAGATGG 755
Db      3093680 GCGCCGCGGTGTCAACGATCAACGCGG 3093651
```

RESULT 4

```
US-08-916-421B-1/c
/ Sequence 1, Application US/08916421B
/ Patent No. 6503729
/ GENERAL INFORMATION:
/ APPLICANT: Bult et al.
/ TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
/ Patent No. 6503729
/ TITLE OF INVENTION: jannaschii
/ FILE REFERENCE: PB275
/ CURRENT APPLICATION NUMBER: US/08/916,421B
/ CURRENT FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschii
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (28222)..(28222)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (28257)..(28258)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (84773)..(84773)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (84808)..(84808)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (84812)..(84812)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (98120)..(98120)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (98159)..(98159)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (98239)..(98239)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (98266)..(98266)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (98343)..(98343)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (103998)..(103998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (148948)..(148948)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (163385)..(163385)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (191989)..(191989)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (191995)..(191995)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980) ..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187) ..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220) ..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814) ..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398) ..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418) ..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837) ..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993) ..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226) ..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167) ..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241) ..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992) ..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708) ..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081) ..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203) ..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435) ..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442) ..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652) ..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684) ..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455) ..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676) ..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539) ..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619) ..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830) ..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
```

```
NAME/KEY: misc feature
LOCATION: (1096846) ..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881) ..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881) ..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988) ..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224) ..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473) ..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491) ..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091) ..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020) ..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912) ..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734) ..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998) ..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854) ..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 9.3%; Score 92; DB 4; Length 1664976;
Best Local Similarity 49.4%; Pred. No. 2.6e-15;
Matches 307; Conservative 0; Mismatches 300; Indels 15; Gaps 2;

```
QY 46 ATATGCAAAATATAGCAAAATTTGACACTTTGCAAAATTTGATGCAAAATCATATA 105
DB 1610456 ATATGTATAGATAGCTTAAGAGAGATATCAATTTGAGAGCTTATCATGATG 1610397

QY 106 AACATTCAGTTAATACCCGATTTCCAGTCCAGATGAATACAGTTAGATGAAGTGA 165
DB 1610396 AAAACGCCCTCAAAAATAGATATGCACAAATTTCCATTTAAGGTTACATTGAGCAGGTTTC 1610337

QY 166 TGGCATCTAAGAAGTTTGTGATTTTGTACTATCTTACAATGTTGTCACCTTCAGATG 225
DB 1610336 TCAACTCTCACCAGTTGCTGAGCCTTTAAGATTACTACATTGCTCACCAGTTTCAGATG 1610277

QY 226 GTGCTGACGAGCAATTTGGCCAGTGAAGCAATTTGTACAGAAGTATGGCCTGCATCCA 285
DB 1610276 GTGCTGCTGCCCTTATAGTATGTGAAGCAGAAAAGGCTTAAAGATTTGTAATTAAGATG 1610217

QY 286 AAGCTGTGAAATTTTGGCAAGAATGATGACTGATTTGCCAAGCTCGTTGAAGAAA 345
DB 1610216 ATATCATCTACATCAAGCAAGTGTCAAGCATCAGATA-----CAATTGCAATTAC 1610166

QY 346 AAGCATTTAATAATGTTGGCTTTGATATGATGAAGAAGCTGCAAGAAAATGCTATG 405
DB 1610165 ACAGCAGAGAGAGCATTAACAGCTTAAAGCTGCTAAAGTTGCAAGTGAAGAAAAGCATATA 1610106

QY 406 AGAAATCTGGCCTGACACCAATATGATATGACGTAATAGAACTTCACGATTCCTTTCTA 465
DB 1610105 AAATGGCAATATAGAACCAAGGATGTGATGTGCGTGAAGTTTCATGACTGCTTTGCTA 1610046
```

QY 466 CCAACGACTCCTTACTTATGAAGCACTGGACTCTGTCCAGAGACAAGTGCAACGC 525
DB 1610045 TAAATGGTTAATATATGATGAGAGAACTGGTTCTGTAAGAGAGAGCGTGAAGA 1609986
QY 526 TGGTTGAT-----AGAGAGATAATACATATGAGAGAAAGTGGTCAATAATCCTAGTG 579
DB 1609985 TAGTTATGATAAAAAAGATAGCTATGATGATGCTCCAGCAGTAATCCAAGTG 1609926
QY 580 GTGACTGATTTCAAGGACACCACTAGCGCTACAGCTCTTCTCAGTGTGAGAAC 639
DB 1609925 GAGGCTGAAAGCTGCTGACATGATTAAGGGCTACCGGTATAGGCAGGTTGAGAGA 1609866
QY 640 TCTGCTGGCAGCTGAGAGGGA 661
DB 1609865 TTACTGGCAGTTAAAGCAGGA 1609844

RESULT 5

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fls
; US-08-232-463-14

Query Match 4.9%; Score 49.2; DB 1; Length 7218;
Best Local Similarity 2.5%; Pred. No. 0.00041;
Matches 9; Conservative 204; Mismatches 153; Indels 0; Gaps 0;
QY 628 AGTGTGAGAACTCTGCTGGCAGCTGAGAGGGAAGCCGAAAGAGCAAGTTCTCTG 687
DB 1453 AGATGAGAAGATTGTGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1394

QY 688 TGCAAAGGTGGCTCTGCNGCATTAATTANGCATTTGAGAGAACTGTGTTACACTCTA 747
DB 1393 RRR 1334
QY 748 CAAGATGGGGTTTCCCGGAAGCCCGCAGTTCCTTTAGAACTCATCAAAATTGAAGCCN 807
DB 1333 RRR 1274
QY 808 GTTCCAACCAAGCTCTGCAAGTATNGGTTTAANGNAAAATCTNGTTTAAAGNGGATT 867
DB 1273 RRR 1214
QY 868 GAGAAAGAAACNTNAAAGAGGGAANGGGAACAATTTGTGAAGAAATAATNCGNGGAA 927
DB 1213 RRR 1154
QY 928 TTTTGCCTTCAAGGGGAANAATGGCCCTTGGGGGTTAAAGANGCCACCTGGGGTG 987
DB 1153 RRR 1094
QY 988 GTGGGA 993
DB 1093 RRRRRR 1088

RESULT 6

US-08-785-048-4
; Sequence 4, Application US/08785048
; Patent No. 5763246
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5763246e1 tRNA Synthetase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,048
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9601099.6
; FILING DATE: 19-JAN-1996
; APPLICATION NUMBER: 9615845.6
; FILING DATE: 27-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimml, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31355-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1644 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-785-048-4

	Query Match	4.0%;	Score 39.8;	DB 1;	Length 1644;	
	Best Local Similarity	48.9%;	Pred. No. 0.1;	Mismatches 112;	Indels 0;	Gaps 0
	Matches 107;	Conservative	0;			
QY	15 CCTCAGATGTTGGGTATGCTGGAAAGAACATATGAAAAATATGAACAATAATTGAA					74
Db	634 CCGAATTAAGATTATTCTGAAGACACGTTTGAAGAATTGAAAATTAGCGTA					693
QY	75 CACTTGCAAAAATTGGATGGAATAATCATAAACATTCAGTTAATAACCCGATTTCCAG					134
Db	694 GATACGAATAGGCTAAATTGAAAAATGATTTAGCAGAGTTCAATACGATTTTGATAAT					753
QY	135 TTCCAAGATGAATACAGTTTAGATGAAGTATGCATCTAAGAAGTTTTGATTTTTTG					194
Db	754 TGGTTTAGTGAACATCTTTATATGAAAAAGCGGAATTCCTGGAAGTTTTAGCAAAATG					813
QY	195 ACTATCTTACAATGTTGTCACCATTGAGATGCTGCA					233
Db	814 AAAGAATTAGGTTATACGTATGAAGCTGATGGCGCTACA					852

```

RESULT 7
US-08-996-799-4
/ Sequence 4, Application US/08996799
/ Patent No. 5965416
/ GENERAL INFORMATION:
/ APPLICANT: Hodgson, John
/ APPLICANT: Lawlor, Elizabeth
/ TITLE OF INVENTION: No. 5965416el tRNA Synthetase
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SmithKline Beecham Corporation
/ STREET: 709 Swedeland Road
/ CITY: King of Prussia
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19406-0939
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/996,799
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/785,048
/ FILING DATE: 17-JAN-1997
/ APPLICATION NUMBER: 9601099.6
/ FILING DATE: 19-JAN-1996
/ APPLICATION NUMBER: 9615845.6
/ FILING DATE: 27-JUL-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Gimm1, Edward R
/ REGISTRATION NUMBER: 38,891
/ REFERENCE/DOCKET NUMBER: P31355-3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 610-270-4478
/ TELEFAX: 610-270-5090
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1644 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ US-08-996-799-4

Query Match      4.0%; Score 39.8; DB 2; Length 1644;
Best Local Similarity 48.9%; Pred. No. 0.1;
Matches 107; Conservative 0; Mismatches 112; Indels 0; Gaps 0

```

[illegible]

```

RESULT 8
US-08-785-048-1
; Sequence 1, Application US/08785048
; Patent No. 5763246
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5763246el tRNA Synthetase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,048
; FILING DATE: 17-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9601099.6
; FILING DATE: 19-JAN-1996
; APPLICATION NUMBER: 9615845.6
; FILING DATE: 27-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm1, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31355-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1662 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-785-048-1

```

	Query Match	4.0%;	Score 39.8;	DB 1;	Length 1662;
	Best Local Similarity	48.9%;	Pred. No. 0.1;		
	Matches 107; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0
OY	15 CCTCAGATGTTGGGTATGCTGGAAGAACAATATGGAACAAATTGAA	74			
Db	652 CCGAATTAAAGATTATTCGAGAAGCACCGTTGAAGAATTAGAAAAATTAGCCGTA	711			

[illegible]

RESULT 9

```

US-08-996-799-1
Sequence 1, Application US/08996799
Patent No. 5965416
GENERAL INFORMATION:
APPLICANT: Hodgson, John
TITLE OF INVENTION: Lawlor, Elizabeth
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/785,048
FILING DATE: 17-JAN-1997
APPLICATION NUMBER: 9601099.6
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 9615845.6
FILING DATE: 27-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31355-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1662 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-996-799-1

```

Query Match	4.0%	Score 39.8;	DB 2;	Length 1662;
Best Local Similarity	48.9%	Pred. No. 0.1;		
Matches 107; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0

Qy	15	CCTCAGATGTTGGGTATGCTGGA	74
Db	652	CCTGAATTTAAAGATTATCTGAA	711
Qy	75	CAC TTGCAAAAATTGATGAAAAAT	134
Db	712	GAATACGAAATGCGTAAATTGAAA	771

QY	135	TTCCAAGATGATAACAGTTTACAGTAGAAGTGATGGCATCTAAAGAAGTTTTGATTTTTTG	194
Db	772	TGGTTTACTGAACATCTTTATATGAAAAAGCGGAATTCTTGAA GTTTTACCAAAAATG	831
QY	195	ACTATCTTACAATGTGTGCCCACTTCAGATGGTGTCGA	233
Db	832	AAAGAATTAGGTTATACGTATGAAGCTGATGGCGCTACA	870

RESULT 10

```

US-08-545-528D-1/c
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

```

```

Query Match          3.9%;   Score 39;   DB 4;   Length 580073;
Best Local Similarity 52.7%;   Pred. No. 2.7;
Matches 107; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

```

QY	9	GTTGCTCCTCAGATGTTTGGGTATGCTGGAAAAAACAATATGGAATAATATGGAACAAA	68
Db	518329	GATGCTCATTGGATGAGTATAAATGAGTTTAAAGAAAAATGAAAAAGATAGAAAAATC	518270
QY	69	ATTGAACACTTTGCAAAAATTGGATGAAAAATCATAA-ACATTCAGTTAATAACCCGTA	127
Db	518269	ATTGAAGCATTTACTAATAATCGTGAATAAATCAACGATATTAAAGCTAGTAATGGA	518210
QY	128	TTCCCAAGTTCCCAAGATGAATACAGTTTAGATGAAGTGATGGCATCTAAAGAGTTTGTGA	187
Db	518209	TTAGTAGTTAGACCAAGAAACTGAAGATAGTAGACGTGAAGCATTCAACAGTAATGGGAT	518150
QY	188	TTTTTTGACTATCTTACAATGTT	210
Db	518149	ATTTTAAACATACACCTTATATT	518127

RESULT 11

```

US-09-328-352-1034
; Sequence 1034, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1034
; LENGTH: 1368
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1034

```

```
Query Match          3.8%; Score 38.2; DB 4; Length 1368;
Best Local Similarity 56.9%; Pred. No. 0.27;
Matches 70; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY      387  GCTGCAGAAATGCTATGAGAATCTGGCCCTGACACCAATGATATTGACGTAATAGAA 446
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       1072 GCTGCACCGCAAGCCTTGGAAGAACTGGTTTAACAATTGATGATTAATGACCTATTGAA 1131

OY      447  CTTACGATTGCTTTTCTTACCAAGCACTCCTTACTTATGAAGCACTGGACTCTGTCCA 506
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       1132 GTTAACGAAGCATTGTGACGCTGTCGATGCGTTTATAAATGAACCTTAATGTACCTGCT 1191

OY      507  GAA 509
          |||
Db       1192 GAA 1194

RESULT 12
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
; Patent No. 6503729
; TITLE OF INVENTION: jannaschi
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschi
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
```

```
OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
```

```
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664854)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1
```

```
Query Match 3.8%; Score 37.8; DB 4; Length 1664976;
Best Local Similarity 52.2%; Pred. No. 9.7;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```
QY 43 AACATATGAAAAATATGGAACAAAAATTGACACTTTGCAAAAAATTGGATGGAATAATC 102
   |||||
Db 732922 AAGATTTTGTGATATATATGATTTTAAAAACCTGAAATAGTTAAGATTAAGAAAAAG 732981
   |||||

QY 103 ATAAACATTCAGTTAATAACCCGTAATTCACAGTCCCAAGATGAATACAGTTAGATGAAG 162
   |||||
Db 732982 AAAATGAAATGAAAAATAGCTTCTCTTAAATCTGAAAAATATGATTATGAT 733041
   |||||

QY 163 TGATGGCATCTAAGAAGTTTGTGATTTTGTGACTATCTTA 203
   |||||
Db 733042 TAGTTGTAGCTCCAGAGATGAAGACATTTATATATTTA 733082
   |||||
```

```
RESULT 13
US-09-252-991A-13040
; Sequence 13040, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13040
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-13040
```

```
Query Match 3.8%; Score 37.4; DB 4; Length 1353;
Best Local Similarity 49.2%; Pred. No. 0.45;
Matches 98; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```
QY 406 AGAATCTGGCCTGACACCAATGATATTGACGTAATAGAACTTCACGATTCCTTTCTA 465
   |||||
Db 1022 AGATGGCCGGTGTGCGGACCGGACATCGACCTGTGATGCTCTACGACGCTTACCA 1081
   |||||

QY 466 CCAACGAACCTCTTACTTATGAAGCACTGGGACTCTGTCCAGAAGACAAAGTGCACGC 525
   |||||
Db 1082 TCAACACCTGCTGTCTCGAGGACCTGGGATTTGCGCCAAAGGCGAGGCGGCCCAT 1141
   |||||

QY 526 TGGTGTATAGAGAGATATATCATATGAGAGAAAGTGGTCTAATCTTAGTGTGAC 585
   |||||
Db 1142 TCGTCCAGGCGGACGATCGCGCCCGGCGGAGTTGGCGGTGAACACCAACGTTGCG 1201
   |||||

QY 586 TGATTTCAAGGACACC 604
   |||||
Db 1202 GGCTGTCTGCTGATCC 1220
   |||||
```

```
RESULT 14
US-09-252-991A-12438/c
; Sequence 12438, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12438
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (93), (101)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
; US-09-252-991A-12438
```

```
Query Match 3.8%; Score 37.4; DB 4; Length 1410;
Best Local Similarity 49.2%; Pred. No. 0.46;
Matches 98; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```
QY 406 AGAATCTGGCCTGACACCAATGATATTGACGTAATAGAACTTCACGATTCCTTTCTA 465
   |||||
Db 461 AGATGGCCGGTGTGCGGACCGGACATCGACCTGTGATGCTCTACGACGCTTACCA 402
   |||||

QY 466 CCAACGAACCTCTTACTTATGAAGCACTGGGACTCTGTCCAGAAGACAAAGTGCACGC 525
   |||||
```

Db 401 TCACACACCTGCTGTTCTCGAGACCTGGATTCTGCGCCAAGGCGAGGGCCCAT 342
QY 526 TGGTGTATAGAGAGATATACATATGAGAAAGTGGTCATAAATCCTAGTGTGAC 585
Db 341 TCGTCCAGGCGGACGCATCGCCCCGGCGGAGTTGGCGGTGAACACCAACGATGCC 282

Db 3538 GATAATTATTTTTTTTAAATATATATT 3568
Search completed: November 27, 2003, 12:32:29
Job time : 85.3681 secs

QY 586 TGATTTCAAGGACACCC 604
Db 281 GGCTGCTCGCTGCATCC 263

RESULT 15

US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 3.6%; Score 35.8; DB 2; Length 19124;
Best Local Similarity 52.3%; Pred. No. 4.5;
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 52 AAAAAATGAGACAAAAATGAACTTGC AAAAATGATGAGAAATCATAAACATT 111
Db 3418 AAAAAATGAGAAAAAATTCGAATATATAAATATTATAAGAAATAGATATACGATGTC 3477
QY 112 CAGTTAATAACCCGATTTCCAGTTCAGATGATACAGTTAGATGAAGTGTGGCAT 171
Db 3478 GAGCTATTAGCGGTAATTAAAGTATGTGAATTTTCATTTAATATGCTATGATCATTT 3537
QY 172 CTAAGAAGTTTGTGATTTTGTACTACTT 202

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:02:42 ; Search time 3414.44 Seconds
(without alignments)
10100.295 Million cell updates/sec

Title: US-09-835-992A-22

Perfect score: 843

Sequence: 1 ggcacaaataaataattttna.....taatccgaacataaaaag 843

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_stb:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_ey:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	725	86.0	843	6	ARI46582	ARI46582 Sequence
2	725	86.0	843	6	BD079831	BD079831 Cancer-as
3	505	59.9	1052	9	HSSCPXT16	U11313 Human stero
4	505	59.9	122176	2	AL358233	AL358233 Homo sapi
5	505	59.9	175046	2	AC022728	AC022728 Homo sapi
6	505	59.9	193774	9	AL445183	AL445183 Human DNA
7	504	59.8	2572	9	HUMSCP2A	M75883 Human stero
8	502.2	59.6	1439	9	BC005911	BC005911 Homo sapi
9	436.2	51.7	714	6	ARI46579	ARI46579 Sequence
10	436.2	51.7	714	6	BD079828	BD079828 Cancer-as
11	411.8	48.8	687	6	ARI46580	ARI46580 Sequence
12	411.8	48.8	687	6	BD079829	BD079829 Cancer-as
13	363.2	43.1	1229	9	S52450	S52450 sterol carr
14	357	42.3	1219	9	HUMSTEEA	M55421 Human stero
15	262.8	31.2	1500	9	HUMSCP2B	M75884 Human stero
16	140	16.6	2661	4	AF051897	AF051897 Oryctolag
17	110	13.0	173	6	AR246194	AR246194 Sequence
18	105.6	12.5	145564	2	AC134794	AC134794 Mus muscu
19	101.8	12.1	2599	10	RAT60KDA	M62763 Rat 60 kDa
20	99.4	11.8	2626	10	BC018384	BC018384 Mus muscu
21	99.4	11.8	248677	2	AL844206	AL844206 Mus muscu
22	96.8	11.5	2571	10	RATSCPXA	M57453 Rat sterol
23	90.6	10.7	904	10	S80339	S80339 SCP-2=stero
24	81.8	9.7	151259	2	AC138822	AC138822 Homo sapi
25	81.8	9.7	160701	9	AC108106	AC108106 Homo sapi
26	81.8	9.7	172404	9	AC138948	AC138948 Homo sapi
27	81.8	9.7	181154	2	AC138842	AC138842 Homo sapi
28	81.8	9.7	185785	9	AC138817	AC138817 Homo sapi
29	81.8	9.7	188885	9	AC138850	AC138850 Homo sapi
30	81.8	9.7	192925	2	AC138953	AC138953 Homo sapi
31	81.8	9.7	201419	9	AC138908	AC138908 Homo sapi
32	81.8	9.7	211297	2	AC138860	AC138860 Homo sapi
33	81.8	9.7	216200	2	AC138862	AC138862 Homo sapi
34	81.8	9.7	225384	9	AC138832	AC138832 Homo sapi
35	81.8	9.7	238249	2	AC138960	AC138960 Homo sapi
36	80.2	9.5	130416	9	AC008851	AC008851 Homo sapi
37	69.4	8.2	263	6	A74403	A74403 Sequence 89
38	69.4	8.2	263	6	A77382	A77382 Sequence 89
39	67.4	8.0	237844	2	AC098628	AC098628 Rattus no
40	67.4	8.0	240155	2	AC129826	AC129826 Rattus no
41	61.8	7.3	18133	6	AX345843	AX345843 Sequence
42	61.8	7.3	18133	6	AX348405	AX348405 Sequence
43	57.2	6.8	1945	10	RNNSLIPTR	X60654 Rat mRNA fo
44	56	6.6	791	10	RATSCP2A	M57454 Rat sterol
45	55.2	6.5	1980	10	BC034613	BC034613 Mus muscu

ALIGNMENTS

RESULT 1
ARI46582
LOCUS ARI46582 843 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 22 from patent US 6218521.
ACCESSION ARI46582
VERSION ARI46582.1 GI:15109771
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 843)
AUTHORS Obata,Y.
TITLE Isolated nucleic acid molecules associated with gastric cancer and
JOURNAL Patent: US 6218521-A 22 17-APR-2001;

Pred. No. is the number of results predicted by chance to have a

FEATURES		Location/Qualifiers	
source		1..843	
BASE COUNT		237 a 184 c 89 g 215 t 118 others	
ORIGIN			
Query Match			
Best Local Similarity 100.0%; Score 725; DB 6; Length 843;			
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GGCCAAAAAANTTATTNAATTTCTATTAAACNTCTCCNCAAAACATTATTNACCC	60
Db	1	GGCCAAAAAANTTATTNAATTTCTATTAAACNTCTCCNCAAAACATTATTNACCC	60
QY	61	TATNNCNCNGANTTTNANAANTAACCTTNTNTTAAAAAACCCTNGAAAAAAATAA	120
Db	61	TATNNCNCNGANTTTNANAANTAACCTTNTNTTAAAAAACCCTNGAAAAAAATAA	120
QY	121	TNGCAATANTTAACCTTNCCTGAAAANGGAATTTNTACCAANGACNGAANCNTNT	180
Db	121	TNGCAATANTTAACCTTNCCTGAAAANGGAATTTNTACCAANGACNGAANCNTNT	180
QY	181	AATTNGAANTNAAATTATANTNGAAGGCGCNCNGAAGAACCAANTCTNATGTCCAATT	240
Db	181	AATTNGAANTNAAATTATANTNGAAGGCGCNCNGAAGAACCAANTCTNATGTCCAATT	240
QY	241	ATCCTNAANGAGGNNTTNANNACTAATNCCCGATTTTCCAATANGAAGAACCCNNNT	300
Db	241	ATCCTNAANGAGGNNTTNANNACTAATNCCCGATTTTCCAATANGAAGAACCCNNNT	300
QY	301	AAAANTNTTTTNAATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAAT	360
Db	301	AAAANTNTTTTNAATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAAT	360
QY	361	TGGAFTGGGAAAAAATNCNGTCCNNATACCCNNGAANNCAANTTTTAAATTTTAA	420
Db	361	TGGAFTGGGAAAAAATNCNGTCCNNATACCCNNGAANNCAANTTTTAAATTTTAA	420
QY	421	ACCCCCCTANTTTTAAANCTATNGAAGAAANTNGATTANNGACTTGAATTGCCAACCCCTAN	480
Db	421	ACCCCCCTANTTTTAAANCTATNGAAGAAANTNGATTANNGACTTGAATTGCCAACCCCTAN	480
QY	481	TTNCNGGCCACCCNGTGGCGNTNGTTCCTTACTTANTCCCCCAAGGAAANNCCCTTAAN	540
Db	481	TTNCNGGCCACCCNGTGGCGNTNGTTCCTTACTTANTCCCCCAAGGAAANNCCCTTAAN	540
QY	541	CNGAANCTCCNCAAAATAACCTTAANTATCCTTGTAAACCAANCAAAACCTTTTNG	600
Db	541	CNGAANCTCCNCAAAATAACCTTAANTATCCTTGTAAACCAANCAAAACCTTTTNG	600
QY	601	TTTACNTANTCCTTGGATTAAAGGGTCCCAATTTNATCCNGAACCCANTTTTCCCCC	660
Db	601	TTTACNTANTCCTTGGATTAAAGGGTCCCAATTTNATCCNGAACCCANTTTTCCCCC	660
QY	661	NAACCATANTTACCATTTTACCTTGGTAAGGCNCAGTNGTTGCANTNCCGAAANCAGT	720
Db	661	NAACCATANTTACCATTTTACCTTGGTAAGGCNCAGTNGTTGCANTNCCGAAANCAGT	720
QY	721	ANTNTTCCCCNGGCNTTTCCTCCCGANCTTGGGAAAAACGGGATNGGTCCCCCCTTAA	780
Db	721	ANTNTTCCCCNGGCNTTTCCTCCCGANCTTGGGAAAAACGGGATNGGTCCCCCCTTAA	780
QY	781	AAAACAACCTTCCCCCNCTTTGGGCCAGGNNTTNTTCCCGTCTAAATCCGAACAATATA	840
Db	781	AAAACAACCTTCCCCCNCTTTGGGCCAGGNNTTNTTCCCGTCTAAATCCGAACAATATA	840
QY	841	AAG 843	
Db	841	AAG 843	

RESULT 2

DD079831

RESULT 2
BD079831

LOCUS	BD079831	843 bp	DNA	linear	PAT 27-AUG-2000
DEFINITION	Cancer-associated nucleic acids and polypeptides.				
ACCESSION	BD079831				
VERSION	BD079831.1	GI:22625434			
KEYWORDS	JP 2001516009-A/497.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I., Oghare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.				
TITLE	Cancer-associated nucleic acids and polypeptides				
JOURNAL	Patent: JP 2001516009-A 497 25-SEP-2001;				
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH				
OS	Homo sapiens (human)				
PN	JP 2001516009-A/497				
PD	25-SEP-2001				
PF	15-JUL-1998 JP 2000503425				
PR	17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR				
10-OCT-1997 US	60/061765,10-OCT-1997 US 08/948705 PR				
11-OCT-1997 GB	9721697.2,22-JUN-1998 US 09/102322 PI LLOYD				
J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI	CHEN, IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI				
PI	OZLEM TURECI,				
PI	UGUR SAHIN				
PC	G01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC				
A61P35/00,	PC C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,				
PC	C12N15/00				
CC	Cancer-associated nucleic acids and polypeptides. FH Key				
location/Qualifiers	1..843				
FT	source				
FT	location/Qualifiers				
1..843	/organism='Homo sapiens (human)'				
/organism="Homo sapiens"					
/mol_type="genomic DNA"					
/db_xref="taxon:9606"					
BASE COUNT	237 a	184 c	89 g	215 t	118 others
ORIGIN					
Query Match	86.0%; Score 725; DB 6; Length 843;				
Best Local Similarity	100.0%; Pred. No. 7.6e-139;				
Matches 843; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	GGCCAAAAAANTTATTNAATTTCTATTAAACNTCTCCNCAAAACATTATTNAACC	60		
Db	1	GGCCAAAAAANTTATTNAATTTCTATTAAACNTCTCCNCAAAACATTATTNAACC	60		
QY	61	TATNNCNCNGANTTTNANAANTAACCTTNTNTTAAAAAAACCTNGAAAAAAATAA	120		
Db	61	TATNNCNCNGANTTTNANAANTAACCTTNTNTTAAAAAAACCTNGAAAAAAATAA	120		
QY	121	TNGCAATANTTAACCTTNCCTGAAAANGGAATTTNTACCAANGACNGAANCNTNT	180		
Db	121	TNGCAATANTTAACCTTNCCTTGAAGAANGAATTTNTACCAANGACNGAANCNTNT	180		
QY	181	AATTNGAANTNAAATTATANTNGAAGGCGCNCNGAAGAACCAANTCTNATGTCCAATT	240		
Db	181	AATTNGAANTNAAATTATANTNGAAGGCGCNCNGAAGAACCAANTCTNATGTCCAATT	240		
QY	241	ATCCTNAANGAGGNNTTNANNACTAATNCCCGATTTTCCAATANGAAGAACCCNNNTT	300		
Db	241	ATCCTNAANGAGGNNTTNANNACTAATNCCCGATTTTCCAATANGAAGAACCCNNNTT	300		
QY	301	AAAANTTTTNAATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAAT	360		
Db	301	AAAANTTTTNAATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAAT	360		
QY	361	TGGAATGGGAAAAAATNCNGTTCNNATACCCNNGAANNCAANTTTTAAATTTTAA	420		

Db 668 TTGATTTGGAGAAATACTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTTA 609
QY 421 ACCCCCTANTTTTAAANCTATNGAAAANTNGATTANNGACTGAATTGC-CAACCCTA 479
Db 608 ATCTCACTAATTTTAAGAACTATGAGAAATTGATTAAATGACATGAAGTGACACAACACTA 549
QY 480 NNTNCGGCCACCCNGTGGCNCNTGNTTCTTACTTANTCCCCCAAGGAAANNCTTAA 539
Db 548 ATTAAGTGGCCAGCTGTTGCAATGTGTTCTTACTAGTTCCTCCCAAGGAAACTCTTAA 489
QY 540 NCGAANCTCCNCCAAATAACCTTAANTATCCTTGTAACCAANCAAAACCTTTTN 599
Db 488 ACTGAATCTCAGCAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTT 429
QY 600 GTTACNTANTCCTTGGAATTAACGGGTCCCAATTNATCCNGAACCCANTTTCCCC 659
Db 428 GTTTACATAGTCTTGGATTTTACTGTCTTAATTTATTTCTGAAACTCAATTTTACCC 369
QY 660 CNAACCATANTTACCATTTTACCTTGTAAGCNCAGTNGTTGCANTNCCGCAANACAG 719
Db 368 CAGACCATTAATTACCATATTAACTTTGTAATGACAGATTGTAATGCAATTCCGCAAGCAG 309
QY 720 TANTNTTCC--CCNGCNCCTTTCCCGCCGANCCTTGGAAAAACGGGATNGTCCCCCT 777
Db 308 TAGTATACCATCAGGCTCTATTCAACCCAGACATAGGAAAAACAGATTGTCACACCCA 249
QY 778 T--AAAAACAACCTTCCCGCCTTTGGCCAGCAGNNTTNTCCCGTCTAATCCGAAC 834
Db 248 TTAAGAAACAATCTCTATCCACGCTATTGGCGATGTGATGTTCACTAATCTGAC 189
QY 835 AATAAAAAG 843
Db 188 AATAAAAATG 180

RESULT 4
AL358233 122176 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 1 clone RP5-835A17, *** SEQUENCING IN
DEFINITION PROGRESS ***, 8 unordered pieces.
ACCESSION AL358233 GI:9797852
VERSION HTG; HTGS PHASE1; HTGS_CANCELLED.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS McLay, K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Aug 12, 2000 this sequence version replaced gi:9214208.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: djb35A17
----- Summary Statistics
Sequencing vector: XGAP4; version 4.5
Chemistry: Dye-terminator ET-amersham; 20% of reads Chemistry:
Dye-terminator Big Dye; 79% of reads
Consensus quality: 118880 bases at least Q40
Consensus quality: 120216 bases at least Q30
Consensus quality: 120979 bases at least Q20
Insert size: 121476; sum-of-contigs
Insert size: 136681; 7.7% error; agarose-fp
Quality coverage: 4.05x in Q20 bases; sum-of-contigs Quality

coverage: 3.75x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 13486: contig of 13486 bp in length
* 13487 13586: gap of 100 bp
* 13587 30893: contig of 17307 bp in length
* 30894 30993: gap of 100 bp
* 30994 37022: contig of 6029 bp in length
* 37023 37122: gap of 100 bp
* 37123 45381: contig of 8259 bp in length
* 45382 45481: gap of 100 bp
* 45482 56788: contig of 11307 bp in length
* 56789 56888: gap of 100 bp
* 56889 84391: contig of 27503 bp in length
* 84392 84491: gap of 100 bp
* 84492 117843: contig of 33352 bp in length
* 117844 117943: gap of 100 bp
* 117944 122176: contig of 4233 bp in length.
FEATURES
source
1. 122176
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone_1fb="RP5-835A17"
/clone_1lb="RPC1-5"
1. 13486
/note="assembly fragment:01065
fragment_chain:1"
misc_feature 13587..30893
/note="assembly fragment:00181
fragment_chain:1"
misc_feature 30994..37022
/note="assembly fragment:00153
fragment_chain:1"
misc_feature 37123..45381
/note="assembly fragment:00580
fragment_chain:2"
misc_feature 45482..56788
/note="assembly fragment:00948
fragment_chain:2"
misc_feature 56889..84391
/note="assembly fragment:00192"
misc_feature 84492..117843
/note="assembly fragment:00299.0"
misc_feature 117944..122176
/note="assembly fragment:00926"
BASE COUNT 31142 a 30481 c 29651 g 30201 t 701 others
ORIGIN

Query Match 59.9%; Score 505; DB 2; Length 122176;
Best Local Similarity 71.8%; Pred. No. 5.3e-94;
Matches 610; Conservative 0; Mismatches 233; Indels 6; Gaps 3;
QY 1 GGCCAAAAAANTTATTTNAATTTCTATTAANCNTCCNCAANCAATTATTTNACC 60
Db 47280 GGCCAGAAAAAGTATTTAATTTCTATTAACAATTTCTCTCAAGCATTTATTC 47339
QY 61 TATNCGNCNGANTTNANAAANTACCTTTNNNTTAAAAAACCTNGGAAAAAATAA 120
Db 47340 TATATCTCACTGAATTTTAAGAAATACATTAGATTAGAAAACTAGGAAAAAGATTA 47399
QY 121 TNGCAATANTTAACCTTCTGAAAANGAAATTTNTAACANGACNGAANCNTTNT 180
Db 47400 ATGCAGATTAATTAACCTTACATGAAAAAGAAATTAACAAGGAGCTGAGACGTTAT 47459

QY	181	AATTNGAANTAAATTATANTTNGAANCGGCNNCGAAACCAANCTTNATGTCCAATT	240
Db	47460	AAATGAAATGAGATTATTAATTTGAAAACCTGCATCTGAAAGCAAACTTATTGTTCAATT	47519
QY	241	ATCCTNAANGAGGNNTTTNNANNACTAAATNCCNGATTTTCCAAATANGAANCCNNNTT	300
Db	47520	ATTCTTAATGATGGTGTATTATGACTAATACACTGATTTTTCATTAAGGAAACCATGTT	47579
QY	301	AAAANTNTTNTNATTTTAAATAACCCNGTNTCCAACCCCGNATCANATTCTTTNATT	360
Db	47580	AAAAATATTTTATTTTAAATAAGCCGTGTTCAGCTCTGATCATATTTCTTTTATT	47639
QY	361	TGGATTGGGGAATAATNCNGTCCNNATACCNNGAANNGCAANTTTTAAATTTTA	420
Db	47640	TTGATTTGGGAGAAATACTGTTTCTGATGACATGAATGCAAAATTTTATGATTTTAA	47699
QY	421	ACCCCTTANTTTTAAANCTATNGAAAAANTNGATTANNGACTGAATTGC - CAACCTTA	479
Db	47700	ATCTCACTAATTTTAAAGAACTATTGAGAAATGATTAAATGACATGAAGTGACACAACTTA	47759
QY	480	NTTNCNGGCCACNGTGGGCGTNGTNTCTTACTTANTCCCCCAAGAAANNCTTAA	539
Db	47760	ATTACTGGCCAGCTGTGGCATGTGTTCTTACTTAGTTCTCCCAAGAAACCTCTTA	47819
QY	540	NCNGAANCTCNCCAAAATAACCTTAANTATCCTTGSTAACCAANCAAAACCTTTTN	599
Db	47820	ACTGAATCTTCAGCAGAAATAATCCTTAATATACTTTGTGAAGCAAAACAAAGCTTTT	47879
QY	600	GTTTACNTANTCCTTGGGATTTAACGGGTCCCAATTNATCCNGAACCCANTTTTCCCC	659
Db	47880	GTTTACATAGTCTTTGGATTTTACTGTCTTAATTTATTTCTGAAACTCAATTTTACCC	47939
QY	660	CNAACCATANTTACCATTTTACCTTGTGAAGCNCAGTNGTTCANTNCCGCAANCAAG	719
Db	47940	CAGACCATTAATTAACCATATTAACCTTGTATGACACAGTGTATGCAATTCGCAAGCAG	47999
QY	720	TANTNTTCC - CGNGGNCCTTCCCCCGANCTTGGAAAAACGGGATNGTCCCCCCTT	777
Db	48000	TAGTATACCATCAGGCTCTATTTCACCCAGACATAGGAAAAACAGATTGGTCACACCCA	48059
QY	778	T - - - AAAAAACAACCTTCCCCCNCCTTTGGCCCAAGNNTTNTTCCCGTCTAATCCGAAC	834
Db	48060	TTAAGAAACAATACTATCCACGCTATTGGCCAATGTGTAGTTCAAGTCTAATTTCTGAC	48119
QY	835	AATAAAAAG 843	
Db	48120	AATAAAAAG 48128	

RESULT 5	AC022728/c	LOCUS	DEFINITION	AC022728	175046 bp	DNA	linear	HTG 24-AUG-2002
AC022728/c								
LOCUS								
DEFINITION								
AC022728								
SEQUENCE								
34 unordered pieces.								
AC022728								
AC022728.4								
GI:7249198								
HTG; HTGS_PHASE1; HTGS_DRAFT.								
Homo sapiens (human)								
Homo sapiens								
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.								
1 (bases 1 to 175046)								
Birren,B., Linton,L., Nusbbaum,C. and Lander,E.								
Homo sapiens chromosome 1, clone RP11-310J14								
Unpublished								
2 (bases 1 to 175046)								
Birren,B., Linton,L., Nusbbaum,C., Lander,E., Abraham,H., Allen,N.,								
Anderson,S., Baldwin,J., Barna,N., Beckertly,R., Beda,F.,								
Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Casale,A.,								
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,								
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,								
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,								
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,								

TITLE	Direct Submission
JOURNAL	Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 175046)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,

Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczký, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marguis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Olivier, T.M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 175046)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczký, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marguis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mhova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE	Direct Submission
JOURNAL	Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Mar 16, 2000 this sequence version replaced gi:6980310.

ALL repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information -----

Center project name: L5491

Center clone name: 310 J 14

----- Summary Statistics -----

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrad: version 0.960731

consensus quality: 156135 passes at least 040

consensus quality: 165373 bases at 1eart 030

Consensus Quality: 168980 bases at 1e8x
Consensus: 105575 bases at 1e8x

Insert size: 185000: Adarose-fn
 combenbus quality: 1000000 at 1000000000

```
Insert size: 171746; sum-of-conf's
Insert size: 183000; agalose-IP
```

Quality coverage: 3 A in Q20 bases: anarose-

Quality coverage: 3.4 in Q20 bases; avg-of-

quality coverage: 3.6 in Q20 bases; sum-of-or-conflicts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

NOTE: This is a 'working draft' sequence. It c

consists of 34 configs. The true order of the

is not known and their order in this sequence

arbitrary. Gaps between the contigs are represented

uns of N , but the exact sizes of the gaps are

This record will be updated with the finished

as soon as it is available and the accession n

e preserved.

1 56: contig of 56 bp in length

57 156: gap of 100 bp

157 1191: contig of 1035 bp in length

1192 1291: gap of 100 bp
1292 2432: contig of 1141 bp in length
2433 2532: gap of 100 bp
2533 3742: contig of 1210 bp in length
3743 3842: gap of 100 bp
3843 4921: contig of 1079 bp in length
4922 5021: gap of 100 bp
5022 6073: contig of 1052 bp in length
6074 6173: gap of 100 bp
6174 7552: contig of 1379 bp in length
7553 7652: gap of 100 bp
7653 9001: contig of 1349 bp in length
9002 9101: gap of 100 bp
9102 10622: contig of 1521 bp in length
10623 10722: gap of 100 bp
10723 11836: contig of 1114 bp in length
11837 11936: gap of 100 bp
11937 13398: contig of 1462 bp in length
13399 13498: gap of 100 bp
13499 14949: contig of 1451 bp in length
14950 15049: gap of 100 bp
15050 16260: contig of 1211 bp in length
16261 16361: gap of 100 bp
16361 18265: contig of 1905 bp in length
18266 18365: gap of 100 bp
18366 19916: contig of 1551 bp in length
19917 20016: gap of 100 bp
20017 22939: contig of 2923 bp in length
22940 23039: gap of 100 bp
23040 26445: contig of 3406 bp in length
26446 26545: gap of 100 bp
26546 29327: contig of 2782 bp in length
29328 29427: gap of 100 bp
29428 32200: contig of 2773 bp in length
32201 32300: gap of 100 bp
32301 36738: contig of 4438 bp in length
36739 36838: gap of 100 bp
36839 43447: contig of 6609 bp in length
43448 43547: gap of 100 bp
43548 48525: contig of 4978 bp in length
48526 48625: gap of 100 bp
48626 52791: contig of 4166 bp in length
52792 52891: gap of 100 bp
52892 58391: contig of 5500 bp in length
58392 58491: gap of 100 bp
58492 64869: contig of 6378 bp in length
64870 64969: gap of 100 bp
64970 73369: contig of 8400 bp in length
73370 73469: gap of 100 bp
73470 81101: contig of 7632 bp in length
81102 81201: gap of 100 bp
81202 87754: contig of 6553 bp in length
87755 87854: gap of 100 bp
87855 95534: contig of 7680 bp in length
95535 95634: gap of 100 bp
95635 102797: contig of 7163 bp in length
102798 102897: gap of 100 bp
102898 111817: contig of 8920 bp in length
111818 111917: gap of 100 bp
111918 121160: contig of 9243 bp in length
121161 121260: gap of 100 bp
121261 140537: contig of 19277 bp in length
140538 140637: gap of 100 bp
140638 175046: contig of 34409 bp in length.

FEATURES

source
1. 175046
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-310J14"
/clone_lib="RPCT-11 Human Male BAC"

misc_feature 1..56
/note="assembly_fragment
clone_end:T7
vector_side:right"
misc_feature 157..1191
/note="assembly_fragment"
1292..2432
/note="assembly_fragment"
2533..3742
/note="assembly_fragment"
3843..4921
/note="assembly_fragment"
5022..6073
/note="assembly_fragment"
misc_feature 5022..6073
/note="assembly_fragment
clone_end:SP6
vector_side:left"
misc_feature 6174..7552
/note="assembly_fragment"
7653..9001
/note="assembly_fragment"
9102..10622
/note="assembly_fragment"
10723..11836
/note="assembly_fragment"
11937..13398
/note="assembly_fragment"
13499..14949
/note="assembly_fragment"
15050..16260
/note="assembly_fragment"
16361..18265
/note="assembly_fragment"

Query Match 59.9%; Score 505; DB 2; Length 175046;
Best Local Similarity 71.8%; Pred. No. 5e-94;
Matches 610; Conservative 0; Mismatches 233; Indels 6; Gaps 3;

QY 1 GGGCAAAAAAANTTATTNAAATTCCTATTAAACNCTCCNCAAAACATTATTNACC 60
Db 43358 GGGCAGAAAAAGTTATTATTTCTATTAAACATTCTCTCAAGCATTTATTATCC 43299
QY 61 TATNNCNCNGANTTNANAAANTACCTTNNNTTAAAAAACCCTNGAAAAAATAA 120
Db 43298 TATATCTCACTGAATTTAAGAAATTAACATTAGTATTAGAAAAAAGTAA 43239
QY 121 TNGCAATANTTAACCTTCTGAAAAANGAAATTTNACCAANGAACNCAANCTNT 180
Db 43238 ATGCAGATAATTAACCTTACATGAAAAAGAAATTAACAAGAGCTGAGACGTAT 43179
QY 181 AATTNGAANTNAAATTAATTNGAANGCGGCGNCGAACAACCAACTTNATGTCAT 240
Db 43178 AATTGAATGAGATTATAATTGAAACTGCATCTGAAGCAACTTATGTCAT 43119
QY 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNT 300
Db 43118 ATTCTTAATGATGGTGTGTTTATGACTAATACATGATTTTCAATAAGAAACCATG 43059
QY 301 AAAANTNTTTNATTTTAAAAATAACCCNGTTCGAACCCCGATCANATTCCTTNA 360
Db 43058 AAAAATATTTTATTTTAAAAATAAGCCGTGTGTTCAAGCTGATCATATTTCTTA 42999
QY 361 TGGATTGGGAAAAAATNCGTCCNNATACCNNGAANGCAAAANTTTTAAATTTT 420
Db 42998 TTGATTTGGGAAAGAAATACTGTTCTGATAGCATGAATGCAAAATTTTGAATTT 42939
QY 421 ACCCCCTANTTTTAAANCTATNGAANAANTNGATTANNGACTTGAATTGC-CAAC 479
Db 42938 ATCTCACTAATTTTAAAGACTATTGAGAAATGATTAATGACATGAGTGCACAC 42879
QY 480 NTTNCGGCCACCCNGTGGCCTNTGTTCCCTTACTTANTCCCCCAAGGAANNCTTAA 539
Db 42878 ATTACTGGCCAGCTGTGGCATGTGTTTCTTACTTAGTTCTCCCAAGGAAACTCT 42819

QY	540	NCNGAANCCTCCNCCAAAATTAACCCCTTAANTATCCCTTGGTAACCAANCAAAACCTTTT	599
Db	42818	ACTGAATCTTCAGCAGAAATAATCCCTTAATAATACTTTGTAAGCAAAACAAAAGCTTTT	42759
QY	600	GTTTACNTAATCCTTGGGATTTAACGGGTCGCCAATTNATCGNAAACCANTTTTCCC	659
Db	42758	GTTTACATAGTTCTTTGGATTTTACTGTTCCTAATTTATTCGAAACTCAATTTTACC	42699
QY	660	CNAACCATANTTACCATTTTTACCCTTGGTAAGGCNCAGTNGTTGCANTNCCGCAANCA	719
Db	42698	CAGACCATTAATTACCATATTAACTTTGTAATGCACAGTTGTATGCAATTCGCAAAAGC	42639
QY	720	TANTNTTCC--CCNGGCNCTTCCCGCANCTTGGAAAAACGGGATNGTCCCCCT	777
Db	42638	TAGTATACCATCAGGCTCTATTTCACCCAGACATAGAAAAACAGGATTGTCACACCA	42579
QY	778	T----AAAAACAACCTTCCCCCNCTTTGGCCCAAGNNTTNTTCCGCTAAATCCGAAC	834
Db	42578	TTAAGAAACAATCTATCCCAAGCTATTGGCCAAATGTGTAGTTTCAGTCTAAATTCGAC	42519
QY	835	AATAAAAAAG 843	
Db	42518	AATAAAATG 42510	

RESULT 6
 AL445183/c
 LOCUS
 DEFINITION Human DNA sequence from clone RPl1-334n14 on chromosome 1, complete sequence.
 AL445183
 ACCESSION AL445183.19 GI:20068427
 VERSION
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 WALLIS, J.
 REFERENCE Direct Submission
 AUTHORS Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
 TITLE Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 JOURNAL humquerry@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
 COMMENT On Apr7, 2002 this sequence version replaced gi:17939714.

Source

```
Location/Qualifiers
1. .19374
/organism="Homo sapiens"
```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-334A14"
/clone_lib="RPC1-11.2"
BASE COUNT      50815 a 45351 c 44031 g 53577 t
ORIGIN

```

Query Match	59.9%;	Score 505;	DB 9;	Length 193774;
Best Local Similarity	71.8%;	Pred. No. 5e-94;		
Matches 610; Conservative	0;	Mismatches 233;	Indels 6;	Gaps 3;

QY	1	GGCCAAAAAANTTATTTTNAATTTCTATTAANCNTCCCTCCNCAAAACATTAATTTNACC	60
Db	98225	GGCCAGAAAAAGTTATTTTAAATTTCTATTAACATTTCTTCCAAAGCATTAATTTTATCC	98166
QY	61	TATNCGNCNGANGANTTTNAAAAAATACCTTTNTNTTTAAAAAACCCTNGAAAAAAAATAA	120
Db	98165	TATATCTCACTGAATTTTAAAGAAAATAACATTAGTATTAGAAAAAATAAGAAAAAGATAA	98106
QY	121	TNGCAATANTTTAACCTTNCCTGAAAAANGAAATTTNTACAAANGAACNGAAANCNTNT	180
Db	98105	ATGCAGATAATTAAACTTACATGAAAAAGAAAAATTATACAAAGAACTGAGAACGTAT	98046
QY	181	AATNGAANTNAAATTATANTNGAANAANGCGCNCNGAAACCAANCTTNATGTCCAATT	240
Db	98045	AAATGGAATAGAGATTATATTTGAAGAACTGCATCTGAAGCAAACTTATGTTCATTT	97986
QY	241	ATCCTNANGAGGNNTTTANNACTAATNCCNGATTTTCCAATANGGAANCCNNNTT	300
Db	97985	ATTCTTAATGATGTGTTTTATGACTAATACACTGATTTTCAATAAGGAAACCCATGTT	97926
QY	301	AAAAATNTTTTNAATTTAAAAAATAACCCNGTNTCCAACCCNGATCANATTCCTTNAAT	360
Db	97925	AAAAATATTTTATTTTAAAAAATAAGCCGTGTGTTCAAGCTTGATCATATTTCTTTAAT	97866
QY	361	TGGATTGGGAAAAAAATNCNGTTCNNATACCNNGAANNCAAAATTTTAAATTTTA	420
Db	97865	TTGATTTGGGAAAGAAATACTGTCTGATAGCATGAATGCAAAATTTTAGATTTTAA	97806
QY	421	ACCCCTTANTTTTAAAAAANGTAAGAAAAANTGATTANNGACTGAATTGC-CAACCTTA	479
Db	97805	ATCTCACTAATTTTAAAGAACTATTGAAGAAATGATTATGACATGAAGTGCACAACACTA	97746
QY	480	NTTNCNGCCACCCNGTGGCNGTNTTCTTACTTANTCCCCCAAGAAANNCTTAA	539
Db	97745	ATTACTGGCCAGCTGTGGCATGTGTTTCTTACTTACTTCTCCCAAGAAAACTCTTAA	97686
QY	540	NCNGAANCTCCNCCAAAAATACCCCTTAANTATCCTTGGTAACCAANCAAAACCTTTTN	599
Db	97685	ACTGAATCTTCAGCAGAAATATCTTAATATATACCTTTGTAAAGCAAAAAAAGCTTTT	97626
QY	600	GTTTACNTANTCCTTGGGATTTAACGGGTCCCCCAATTTNATCCNGAAACCCANTTTTCCC	659
Db	97625	GTTTACATAGTCTTTGGATTTTACTGTCTTAATTTTATTTCTGAAACTCAATTTTACC	97566
QY	660	CNAAACATANTTACCATTTTTACCTTGTAGGCNCAGTNGTTGCANTNCCGCAANCAAG	719
Db	97565	CAGACCATATATTACCATATTTAATTGTATATGACAGTGTATGCAATTCGCAAAAGCAG	97506
QY	720	TANTNTTCC--CCNGGCNTTTCCCCCGANCTTGGGAAAAAAGGATNGTCCCCCT	777
Db	97505	TAGTATATCCATCAGGCTCTATTTCAACCCAGAGCATAGAAAAACAGATTGCTACACCCA	97446
QY	778	T---AAAAAACACCTTCCCCCNCCTTTGGCCCCAGGNNTTNTTCCCGTAAATCCGAAC	834
Db	97445	TTTAAAGAAACAAATCTATCTCCACGCTATTTGGCCAATGTGTAGTTTCAAGTCTAAATCTGAC	97386
QY	835	AATAAAAAG 843	
Db	97385	AATAAAAATG 97377	

RESULT 7	LOCUS.	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear	PRI 06-DEC-1993																	
2572 bp	mRNA	linear</																		

BASE COUNT	794 a	439 c	570 g	769 t
ORIGIN	/gene="SCP-X/SCP-2" /evidence=experimental			
Query Match	59.8%; Score 504; DB 9; Length 2572;			
Best Local Similarity	71.8%; Pred. No. 1.5e-93;			
Matches	609;	Conservative	0;	Mismatches 233; Indels 6; Gaps 3;
QY	2	GCCAAAAAANTTATTTNAAATTCCTATTAAACNTCCCTCCNCAAAACATTAATTNACCCT	61	
Db	2571	GCCAGAAAAAGTTATTTTAATTTCTATTAAACATTTCTCTCAAGCATTAATTATCTCT	2512	
QY	62	ATNNCNCNCNGANTTTNANAANAATACCTTTNNNTNTTAAAAAACCTTNGAAAAAATAAT	121	
Db	2511	ATATCTCACTGAATTTTAAAGAAATACATTAAGTATTAAGAAAACTAGAAAAAAGATAAA	2452	
QY	122	NGCAATANTTAACCTTNCCTTGAAGAANGAAATTNTTACCAANGAAGCAAGCAAGCAAGCAAG	181	
Db	2451	TGCAGATAATTAACCTACATGAAAAAGAAATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAG	2392	
QY	182	ATTNGAANTNAAATTATANTTNGAANGCGCNCNGAAGCAACCAACTTNAATGTTCCAATTA	241	
Db	2391	AATTGAATGAGATTATAATTTGAAAACTGCATCTGAAGCAAACTTATTTGTTCAATTA	2332	
QY	242	TCCTNAANGAGGNNTTNANNACTAATNCCNGAATTTTCCCAATANGAANGCAAGCAAGCAAG	301	
Db	2331	TTCTTAATGATGTTGTTTATGACTAATACATGATTTTCAATTAAGGAAGCAAGCAAGCAAG	2272	
QY	302	AAANTNTTTTNAATTTTAAAAATAACCCNGTTCGAACCCCGATCANATTCCTTNAATTT	361	
Db	2271	AAAAATATTTTATTTTAAAAATAAGCCGTGTTCAAGCTCTGATCATATTTCTTTTAATTT	2212	
QY	362	GGATTGGGAAAAAATNCNGTTCNNATACNNGAANGCAAAATTTTAAATTTTAA	421	
Db	2211	TGATTTGGGAAGAAATACTGTTCTGATAGCATGAAGATGCAAAATTTTGAATTTTAA	2152	
QY	422	CCCCCTANTTTTAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCTAN	480	
Db	2151	TCTCACTAATTTTAAAGACTATTGAGAAATGATTAATGACATGAAGTGACACACACTAA	2092	
QY	481	TTNCGGCCACCCNGTGGCNGTNTTCCCTACTTANTCCCCCAAGAAANNCTTAAN	540	
Db	2091	TTACTGGCCAGCTGTGGCAATGTGTTCTTACTTAAGTTCTCCCAAGAAACTCTTAA	2032	
QY	541	CNGAANCTCCNCAAAATAACCTTAANTATCTTGGTTAACCAAAACAAACCTTTTNG	600	
Db	2031	CTGAATCTTCAGCAGATAATCCTTAATAATATCTTTGTAAAGCAAAACAAAGCTTTTGTG	1972	
QY	601	TTTACNTANTCCTTGGGATTTAACGGGTCCCAATTNATCCNGAACCCANTTTTCCCCC	660	
Db	1971	TTTACATAGTTCTTTGGATTTTACTGTCTTAATTTTATTTCTGAAACTCAATTTTACCCC	1912	
QY	661	NAACCATANTTACCAATTTTACCTTGTAGGCNAGTNGTTGCANTNCCGAANCAGT	720	
Db	1911	AGACCATAATTACCATATTAACCTTGTAAATGACAGATTGTATGCAATTCGCAAGACAGT	1852	
QY	721	ANTNTTCC--CCNGGNCCTTCCCGGANCCTTGGGAAAAACGGGATNGTCCCGCCCTT	778	
Db	1851	AGTATACCATCAGGCTCTATTCAACCCAGAGCATAAGAAAAACAGATTGGTCAACCCAT	1792	
QY	779	---AAAAACAACCTTCCCCNCCTTTGGGCCAGGNNTTNTTCCCGTCTAATCCGAACA	835	
Db	1791	TAAAGAAACAATCTATCCACGCTATTGGCCAATGTGTAGTTTCAAGTCTAATTTCTGACA	1732	
QY	836	ATAAAAAG 843		
Db	1731	ATAAAATG 1724		
RESULT 8				
BC005911/c	BC005911			
LOCUS	BC005911	1439 bp	mrna	linear PRI 12-JUL-2001

DEFINITION Homo sapiens, sterol carrier protein 2, clone MGC:14505
IMAGE:4287946, mRNA, complete cds.
ACCESSION BC005911
VERSION BC005911.1 GI:13543502
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1439)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 21 Row: n Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 432978.
FEATURES
source
location/Qualifiers
1..1439
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocustID:6342"
/db_xref="taxon:9606"
/clone="MGC:14505 IMAGE:4287946"
/tissue_type="Brain, primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
74..505
/codon_start=1
/product="sterol carrier protein 2"
/protein_id="AAH05911.1"
/db_xref="GI:13543503"
/translation="MGPEEAAASSFRTQIEAVPTSSASDGFKANLVFKEIEKLEEG
EQFVKIGIGIFAFKVKDGEKGKATWVDKNGKSVLPNSDKADCTITMADSDFLA
LMTGKMNPOSAPFQKUKITGNMGLAMKLNLOLQPGNAKL"
BASE COUNT 490 a 219 c 272 g 458 t
ORIGIN
Query Match 59.6%; Score 502.2; DB 9; Length 1439;
Best Local Similarity 71.9%; Pred. No. 3.8e-93;
Matches 606; Conservative 0; Mismatches 231; Indels 6; Gaps 3;

Db 1226 AATGAGATTATTAATTGAAAACCTGCATCTGAAGCAAACTTATTGTCATATATCTT 1167
Qy 247 AANGAGGNTTANNACTAATNCCNGATTTTCCATANGGAANCCNNNTTAAANT 306
Db 1166 AATGATGCTTTTATGACTAATACACTGATTTTCAATAGGAAACCATGTATAAAT 1107
Qy 307 NTTTTNATTTTAAATAACCCNGTNTCCAACCCCGCATCANATTCCTTNATTTGATT 366
Db 1106 ATTTTATTTTAAATAAGCCTGTGTCAAGCTCTGATATTTCTTTATTTGATT 1047
Qy 367 GGGGAAAAAATNCNGTTCNNATACCNNGAANNCAANTTTTAAATTTTAAACCCC 426
Db 1046 TGGGAAGAAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCA 987
Qy 427 CTANTTTTAAANCTATNGAAAANTNGATTANNGACTTGAATTGC-CAACCTANTNCN 485
Db 986 CTAATTTTAAAGACTATGAGAAATTGATTATGACATGAAGTGACACACTAATTACT 927
Qy 486 GGGCACCNGTGGGCGNTNGTTCCTTACTTANTCCCCCAAGAAANCCTTAANCNAA 545
Db 926 GGCCAGCTGTGGCATTGTGTTCTTACTTACTTCCCAAGAAACTCTTAACCTGA 867
Qy 546 NCTCCNCAAAATAACCTTAANTATCCTTGTGAACCAANCAAAACCTTTTNGTTAC 605
Db 866 TCTTCAGACAGATAATCCTTAATATACTTTGTAGCAAAACAAAGCTTTTGTGTAC 807
Qy 606 NTANTCCTTGGAATTAAGGGTCCCAATTTNATCCNGAACCCANTTTCCCCCNACC 665
Db 806 ATAGTCTCTTGATTTTACTGTCTCTAATTTATTTGAAACCAATTTTACCCAGACC 747
Qy 666 ATANTTACCATTTTACCTTGTGAAGCNCAGTNGTTGCANTNCCGCAANCACTANTNT 725
Db 746 ATAATTACCATATTAATTGTAATGACACAGTTGTATGCAATTCCGCAAGCAGTAGTAT 687
Qy 726 TCC--CCNGGCNTTTCCTCCCGANCTTGGGAAAAACGGGATNGTCCCCCCTT--AA 780
Db 686 ACCATCAGGCTCTATTACCCAGACATAGAAAAACAGATTGTCACACCCATTAGA 627
Qy 781 AAAACAACCTTCCCCCNCCTTTGGCCCGCAGGNTTNTTCCGCTAATTCGAAACAATAAA 840
Db 626 AACCAATCTATCCACGCTATTGTCATGTAGTTTCAATTAATTCTGACAAATTA 567
Qy 841 AAG 843
Db 566 ATG 564
RESULT 9
ARI46579 714 bp DNA linear PAT 08-AUG-2001
LOCUS ARI46579
DEFINITION Sequence 19 from patent US 6218521.
ACCESSION ARI46579
VERSION ARI46579.1 GI:15109768
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 714)
AUTHORS Obata, Y.
TITLE Isolated nucleic acid molecules associated with gastric cancer and
methods for diagnosing and treating gastric cancer
JOURNAL Patent: US 6218521-A 19 17-APR-2001;
FEATURES location/Qualifiers
source 1..714
BASE COUNT 258 a 100 c 92 g 260 t 4 others
ORIGIN
Query Match 51.7%; Score 436.2; DB 6; Length 714;
Best Local Similarity 73.1%; Pred. No. 1.5e-79;
Matches 516; Conservative 0; Mismatches 187; Indels 3; Gaps 3;

QY 2 GCCAAAAAANTTATTNAAATTCCTTAATTAANCNTCCCNCAANCATTATTNACCCT 61
Db 2 GCCAGAAAAAGTATTTTAAATTTCTTAATAAACATTTCTTCAAGCATTATTATCT 61
QY 62 ATNNCNCNGANTTTNAAAAANTACCTTNNNTTAAAAAACCTNGGAAAAAATAAT 121
Db 62 ATATCTCACTGAATTTTAAGAAATACATAGTATTAGAAAACTAGAAAAAGATAAA 121
QY 122 NGCAATANTTAACCTTNCCTGAAAAANGAAATTNTTACCAANGACNGAANCNTNTA 181
Db 122 TGCAGATTAATTAACCTTACATGAAAAAGAAATTATTAACAAGAGACTGAGAACCTTATA 181
QY 182 ATNGAANTNAAATTATANTNGAANGCGCNCNGAAGAACCAANTTNATGCTCCAATTA 241
Db 182 AATGAAATGAGATTATTAATTTGAAAACTGCATCTGAAGCAAACTTTATTGTTCAATTA 241
QY 242 TCCTNANGAGGNNTTNANNACTAATNCCNGATTTTCCAATNANGAANCCNNNTTA 301
Db 242 TNCTTAATGATGCTGTTTATGACTAATACACTGATTTTCAAGAAAGAAACCATGTTA 301
QY 302 AAANTNTTNNATTTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCCTTNATTT 361
Db 302 AAAATATTTTATTTTAAAAATAGCCTGTGTCAAGCTGTGATCATATTTCTTTATTT 361
QY 362 GGATGGGAAAAAATNCGNTGCCNNAATACCNNGAANNCAANTTTTAAATTTTAA 421
Db 362 TGATTTGGGAANAAAAATACCTGTTCTGATGACATGAATGCAAAATTTTAGATTTTAA 421
QY 422 CCCCCCTANTTTTAAAANCATNGAANAANTNGATTANNGACTTGAATTGC -CAACCTTAN 480
Db 422 TCTCACTAATTTTAANAACATATGAGAAATTTGATTAATGACATGAGAGTGACACACTTAA 481
QY 481 TTNCNGGCCACCNGTGGCGNTNGTTCCTTACTTANTCCCCCAA -GGAANNCTTAA 539
Db 482 TTACTGGCCAGCTGTGTGCAATGTGTTCTTACTTAGTTCGCCAAGGAAAAACTTTAA 541
QY 540 NCNGAANTCCNCCAAATAACCTTAANTATCCTTGTGAACCAANCAAAACCTTTTN 599
Db 542 ATTGAATCTTCAGACAATAATCCTTAATAATACTTTGTGAAGCAAAACAAAGCTTTT 601
QY 600 GTTACNTA -NTCCTTGGGATTTAAGGGTCCCAATTTNATCCNGAACCCANTTTTCC 658
Db 602 GTTACATAGTCTTGGGATTTTACTGTCTCTAATTTTATCTGAAACTCAATTTTACC 661
QY 659 CCNAACCATANTTACCATTTTACCTTGTGTAAGCNCAGTNGTTGC 704
Db 662 CCAGACCATATTAATCAATATTAACCTTGTNTGACAGTGTGTTGC 707

RESULT 10
BD079828 714 bp DNA linear PAT 27-AUG-2002
LOCUS Cancer-associated nucleic acids and polypeptides.
DEFINITION BD079828
ACCESSION BD079828.1 GI:22625431
VERSION JP 2001516009-A/494.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 714)
AUTHORS Old, L.J., Scanlan, M.J., Stockert, E., Gure, A., Chen, Y.T., Gout, I.,
Oghare, M., Obata, Y., Pfeundschuh, M., Tureci, O. and Sahin, U.
TITLE Cancer-associated nucleic acids and polypeptides
JOURNAL Patent: JP 2001516009-A 494 25-SEP-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH
COMMENT OS Homo sapiens (human)
PN JP 2001516009-A/494
PD 25-SEP-2001
PF 15-JUL-1998 JP 2000503425
PR 17-JUL-1997 US 08/896164, 10-OCT-1997 US 60/061599 PR
10-OCT-1997 US 60/061765, 10-OCT-1997 US 08/948705 PR
11-OCT-1997 GB 9721697.2, 22-JUN-1998 US 09/102322 PI LLOYD

J OLD, MATTHEW J SCANLAN, ELISABETH STOCKERT, ALI GURE, YAO PI TSENG
CHEN, IVAN GOUT, MICHAEL O'HARE, YUICHI OBATA, MICHAEL PFEUNDSCUH, PI
OZLEM TURECI,
PI UGUR SAHIN
PC G01N33/574, A61K38/00, A61K39/395, A61K39/395, A61K45/00, A61K48/00, PC
A61P35/00,
PC C07K14/82, C07K16/32, C12N15/09//C07K16/46, C12P21/08, A61K37/02,
PC C12N15/00
CC Cancer-associated nucleic acids and polypeptides. FH Key
location/Qualifiers
FT source 1..714
FT location/Qualifiers
FEATURES
source 1..714
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 258 a 100 c 92 g 260 t 4 others
ORIGIN

QY 2 GCCAAAAAANTTATTNAAATTCCTATTAANCNTCCCNCAANCATTATTNACCCT 61
Db 2 GCCAGAAAAAGTATTTTAAATTTCTTAATAAACATTTCTTCAAGCATTATTATCT 61
QY 62 ATNNCNCNGANTTTNANAANTACCTTNNNTTAAAAAACCTNGGAAAAAATAAT 121
Db 62 ATATCTCACTGAATTTTAAGAAATACATAGTATTAGAAAACTAGAAAAAGATAAA 121
QY 122 NGCAATANTTAACCTTNCCTGAAAAANGAAATTNTTACCAANGACNGAANCNTNTA 181
Db 122 TGCAGATTAATTAACCTTACATGAAAAAGAAATTATTAACAAGAGACTGAGACGTATA 181
QY 182 ATNGAANTNAAATTAANTNGAANGCGCNCNGAAGAACCAANTTNATGCTCCAATTA 241
Db 182 AATGAATGAGATTTAATTTGAAAACTGCATCTGAAGAACTTATTTGTTCAATTA 241
QY 242 TCCTNANGAGGNNTTNANNACTAATNCCNGATTTTCCAATNANGAANCCNNNTTA 301
Db 242 TNCTTAATGATGCTGTTTATGACTAATAATACACTGATTTTCAAGAAAGAAACCATGTTA 301
QY 302 AAANTNTTNNATTTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCCTTTNATTT 361
Db 302 AAAATATTTTATTTTAAAAATAAGCCTGTGTGCAAGCTCTGATCATATTTCTTTATTT 361
QY 362 GGATGGGAAAAAATNCGNTGCCNNAATACCNNGAANNCAAAANTTTTAAATTTTAA 421
Db 362 TGATTTGGGAANAAAAATACCTGTTTCTGATAGCAATGAATGCAAAATTTTGAATTTTAA 421
QY 422 CCCCCCTANTTTTAAAANCATNGAANAANTNGATTANNGACTTGAATTGC -CAACCTTAN 480
Db 422 TCTCACTAATTTTAANAACATATGAGAAATTTGATTAATGACATGAAGTGCAACAACCTAA 481
QY 481 TTNCNGGCCACCNGTGGCGNTNGTTCCTTACTTANTCCCCCAA -GGAANNCTTAA 539
Db 482 TTACTGGCCAGCTGTGTGCAATGTGTTCTTACTTAGTTCCTCCCAAGGAAAACTTTAA 541
QY 540 NCNGAANTCCNCCAAATAACCTTAANTATCCTTGTGTAACCAANCAAAACCTTTTN 599
Db 542 ATGAATCTTCAGACAATAATCCTTAANTATACTTTGTAAGCAAAACAAAGCTTTT 601
QY 600 GTTACNTA -NTCCTTGGGATTTAAGGGTCCCAATTTNATCCNGAACCCANTTTTCC 658
Db 602 GTTACATAGTCTTGGGATTTTACTGTCTCTAATTTTATCTGAAACTCAATTTTACC 661
QY 659 CCNAACCATANTTACCATTTTACCTTGTGTAAGCNCAGTNGTTGC 704
Db 662 CCAGACCATATTAATCAATATTAACCTTGTNTGACAGTGTGTTGC 707

RESULT 11	LOCUS	DEFINITION	SEQUENCE	20 FROM PATENT	US 6218521.	687 bp	DNA	linear	PAT 08-AUG-2001
ARI46580	ARI46580	Sequence 20 from patent US 6218521.							
VERSION	ARI46580								
KEYWORDS	ARI46580.1	GI:15109769							
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	Unclassified.								
AUTHORS	1 (bases 1 to 687)								
TITLE	Obata, Y.								
JOURNAL	Isolated nucleic acid molecules associated with gastric cancer and methods for diagnosing and treating gastric cancer								
FEATURES	Patent: US 6218521-A 20 17-APR-2001;								
Source	Location/Qualifiers								
	1..687								
BASE COUNT	242 a	98 c	85 g	258 t	4 others				
ORIGIN	/organism="unknown"								
Query Match	Best Local Similarity	48.8%	Score	411.8;	DB 6;	Length	687;		
Matches	495;	Conservative	0;	Mismatches	190;	Indels	4;	Gaps	3;
OY	15	ATTNAAATTTCTATTAA	NCNTCCTCCNCA	ANCAATTATTTNA	CCTATNNCNCNGAN	74			
Db	1	ATTTAATTTCTATTAA	CATCTCTCTCA	AAGCATTATTTA	CTCTATATCTCACTGAA	60			
OY	75	TTTNAAAANTACCTT	TNNNTTAA	AAAAAACCCTNG	GA	134			
Db	61	TTTTAANAATAACAT	TAGTATTA	GAAAACTAGG	AAAAAGATNAATGCAGATATTAA	120			
OY	135	CCTTNCCTGAAA	ANGAAATTTN	TACCAANG	GACNGAAANCNTTNAATTNGAANTNAAA	194			
Db	121	ACTTACATGAAAA	AGAAATTA	TAA	CAAGGAGCTGAGAACGTTAATAATTGAATGAGA	180			
OY	195	TTATANTNGAA	ANCGGCNNCGAA	ACCAANTTNAT	GTCCATTAATCTCTNAANGAGG	254			
Db	181	TTATAATTGAAA	AACTGCAT	CGAAAGCA	AACTTATGTTCATTAATCTTAATGATGG	240			
OY	255	NNTTNANNACTA	TATNCCNGAT	TTTCCATANGA	ANCCNNTTAAANTNTTAT	314			
Db	241	TGTTTATGACTA	TATACACTGA	TTTTTCAATA	AGGAAACCCATGTTAAATAATTTTAT	300			
OY	315	TTTAAAAATA	ACCCNGTNTCCA	ACCCNGATCA	NATTCCTTNAATTGGATTGGGAAA	374			
Db	301	TTTAAAAATA	AGCCCTGIGT	CAAGCTCGAT	CATATTTCTTTAATTTTGAATTTGGGAAGA	360			
OY	375	AAATNCNGTTC	CNNNATAC	CNNGAAN	GCAAAANTTTTAAATTTTAAACCCCTANTTTT	434			
Db	361	AAATACTGTT	CTGATAGCA	TGAATGCA	AAAAATTTTAGATTTTAAATCTCNCNTAATTTT	420			
OY	435	AAAAACTATNG	AAAAANTNGA	TTANNGA	CTTGAATTGC-CAACCTANTTNCNGCCACCN	493			
Db	421	AAGAACTATT	GAGAAATG	ATATGACATGA	AGTGACACAACACTAATTA	CTGGCCAGCT 480			
OY	494	GTGGC	NTGNTTCT	CTACTTANT	CCCCCAAGAAANNCTTAA	NCNGAANTCCNCC 553			
Db	481	GTTGGC	ATTGTGTTCT	TACTTACTT	CTCCCAAGGAAACCTCTTAA	ACTGAATCTTCAGC 540			
OY	554	AAAAATA	ACCCCTTA	ANTATCCTT	GGTAACCAAAACCTTTTNGTTACNTANTCCT	613			
Db	541	NGAATA	ACCTTA	ATAATAC	TTTGTTAGCCAAAC-AAAACTTTTGTGTTA	CATAGTTCT 598			
OY	614	TGGGAT	TTAACGGGT	CCCCCAAT	TNATCCNGAACCCANTTTTCCCCNAA	CCATANTTAC 673			
Db	599	TTGGAT	TTTACTGTTCC	TAAATTTA	TTCTGAAACTCCATTTTCCCA	GACCATTAATAC 658			
OY	674	CAT-TT	TACCTTGT	AGGCNCA	GTNGTT 701				

Db	659	CCTATTTAACCTTTGTTATGCACAGTTGTT	687		
RESULT 12					
BD079829					
LOCUS	BD079829	687 bp	DNA linear PAT 27-AUG-2002		
DEFINITION	Cancer-associated nucleic acids and polypeptides.				
ACCESSION	BD079829				
VERSION	BD079829.1	GI:22625432			
KEYWORDS	JP 2001516009-A/495.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 687)				
AUTHORS	Old, L.J., Scanlan, M.J., Stockert, E., Gure, A., Chen, Y.T., Gout, I., Oghare, M., Obata, Y., Pfreundschuh, M., Tureci, O. and Sahln, U.				
TITLE	Cancer-associated nucleic acids and polypeptides				
JOURNAL	Patent: JP 2001516009-A 495 25-SEP-2001;				
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH				
	OS Homo sapiens (human)				
	PN JP 2001516009-A/495				
	PD 25-SEP-2001				
	PF 15-JUL-1998 JP 2000503425				
	PR 17-JUL-1997 US 08/896164, 10-OCT-1997 US 60/061599 PR				
	10-OCT-1997 US 60/061765, 10-OCT-1997 US 08/948705 PR				
	11-OCT-1997 GB 9721697.2, 22-JUN-1998 US 09/102322 PI LLOYD				
	J OLD, MATTHEW J SCANLAN, ELISABETH STOCKERT, ALI GURE, YAO PI TSENG				
	CHEN, IVAN GOUT, MICHAEL O'HARE, YUICHI OBATA, MICHAEL PFREUNDSCUH, PI				
	OZLEM TURECI,				
	PI UGUR SAHIN				
	PC G01N33/574, A61K38/00, A61K39/395, A61K39/395, A61K45/00, A61K48/00, PC				
	A61P35/00,				
	PC C07K14/82, C07K16/32, C12N15/09//C07K16/46, C12P21/08, A61K37/02,				
	PC C12N15/00				
	CC Cancer-associated nucleic acids and polypeptides. FH Key				
	Location/Qualifiers				
	FT source	1.687			
	FT	/organism='Homo sapiens (human)'			
FEATURES	Location/Qualifiers				
source	1.687				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	242 a	98 c	85 g 258 t 4 others		
ORIGIN					
Query Match	48.8%	Score 411.8;	DB 6; Length 687;		
Best Local Similarity	71.8%	Pred. No. 1.5e-74;			
Matches	495;	Conservative 0;	Mismatches 190; Indels 4; Gaps 3;		
QY	15	ATTNAAATTCCTATTAAANCCTCCNCAANCAATTATTNACCCCTATNNCNCNGAN	74		
Db	1	ATTTTAATTTCTTAATAAACATCTCTCTCAAGCATTTATTTATCCTATATCTCACTGAA	60		
QY	75	TTTNANAAANTACCTTNNNTTAAAAAACCCTNGSAAAAAATAATNGCAAAATANTTAA	134		
Db	61	TTTTAANAATAACATTAGTATTAGAAAAAACTAGGAAAAAAGATNAATGCAGTAATTA	120		
QY	135	CCTTNCCTGAAAAANGAAATTTNTACCAANGACNGAAANCNTTNTAATTNGAANTNAA	194		
Db	121	ACTTACATGAAAAAGGAAAAATTATTAACAAAGGAGCTGAGAACGTTATTAATTGAAATGAGA	180		
QY	195	TTATANTTNGAAANGCGCNCNGAAACCAACCTTNAATGGTCCAATTATCCTNAANGAGGG	254		
Db	181	TTATTAATTGAAAACTGCATCTGAAAGCAAACTTATTTGTTCAATTATTTTAATGATGG	240		
QY	255	NNTTNANNACTAATNCCNGATTTTCCAAATANGGAANCCCNNTTAAANNTTTTNAAT	314		
Db	241	TGTTTATGACTAATACACTGATTTTTCATTAAGGAAACCACATGTATAAAATATTTTAT	300		


```
COMMENT      On Dec 6, 1993 this sequence version replaced gi:432972.
FEATURES
  source
    1.1219
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /map="1p32"
      /issue_type="liver"
      1..1219
      /gene="SCP-2"
      /gene="SCP-2"
      84..515
      /gene="SCP-2"
      /codon_start=1
      /product="sterol carrier protein-2"
      /protein_id="AA03559.1"
      /db_xref="GI:432973"
      /translation="MGFPEAASSFRTHQIEAVPTSSASDGFKANLVFKEIEKLDEEG
      EGFVKKIGIFAFKVDGPGKEATWVDVKGKSVLPNSDKADCTITMADSDFLA
      LMTGKNPQSAFPGKLIKITGNMGLMKLQNLQIPGNAKL"
      84..143
      /gene="SCP-2"
      144..512
      /gene="SCP-2"
      /product="unnamed"
      782..787
      /gene="SCP-2"
      /evidence=not_experimental
      polyA_signal
      1062..1067
      /gene="SCP-2"
      /evidence=not_experimental
      polyA_signal
      1107..1112
      /gene="SCP-2"
      /evidence=not_experimental
      polyA_signal
      1190..1195
      /gene="SCP-2"
      /evidence=not_experimental
      polyA_signal
      1219
      /evidence=experimental
      polyA_site
      1219
      /gene="SCP-2"
      /evidence=experimental

BASE COUNT   402 a      195 c      255 g      367 t
ORIGIN
Query Match      42.3%; Score 357; DB 9; Length 1219;
Best Local Similarity 71.4%; Pred. No. 2.5e-63;
Matches 461; Conservative 0; Mismatches 178; Indels 7; Gaps 4;

QY      204 GAAANCGCNCNGAACAACCAANCTTNATGTGCCAATTATCCTNAANGAGGNNTTNANN 263
      |||||
Db      1218 GAAAACTGCATCTGAAGCAAACTTATTGTTCATTATTTCTTAATGATGCTTTATG 1159

QY      264 ACTAATNCCNGATTTTCCATANGAANCCNNNTTAAANTNTTTNATTTTAAAAAT 323
      |||||
Db      1158 ACTAATACACTGATTTTCAATGAAGAAACCATGTTAAAAATATTTTATTTTAAAAAT 1099

QY      324 AACCNGTNTCCAAACCCNGATCANATTCCTTNATTGATGGGAAAAAATNCNGT 383
      |||||
Db      1098 AAGCTGTGTTCAAGCTCTGATCATATTTCTTTATTTGATTGGGAAGAAATACGT 1039

QY      384 TCCNNATACCNNGAANNGAANTTTTAAATTTTAAACCCCTANTTTTAAAAANTAT 443
      |||||
Db      1038 TTCTGATAGCATGAATGCAAAATTTTATGATTTT-ATCTCACTAATTTTAAAGACTAT 980

QY      444 NGAAAAANTNGATTANNGACTTGAATTGC-CAACCTANTTNCGGCCACCGTGGGNTN 502
      |||||
Db      979 TGAGAAATTGATTATGACATGAAGTGCACACACTAATTAATGCGCAGCTGTGGCAAT 920

QY      503 GTNTTCCTTACTTANTCCCCCAAGAAANNCTTAANCNGAANTCCNCCAAAATTAAC 562
      |||||
Db      919 GTGTTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAACCTGAATCTTCAGCAGATTAATC 860

QY      563 CTTAANTATCCTTGTAACCAANCAAAACCTTTTNGTTTACNTANTCCTGGGATTTA 622
      |||||
```

```
Db      859 CTTAATATACTTTGTAAGCAAAACAAAGCCTTTTGTGTTACATAGTCTTGGATTTT 800

QY      623 ACGGTCGCCCAATTTNATCCNGAACCCANTTTTCCCGNAACATANTTACCATTTTACC 682
      |||||
Db      799 ACTGTTCCCTAATTTTATCTGAAGACTCAATTTTACCCAGACCATTAATTACCATATTAC 740

QY      683 TTGCTAAGCNCAGTNGTTTGCANTNCCGCAANACGATANTNTTCC--CCNGCNCCTTTC 740
      |||||
Db      739 TTTGTAATGACACAGTTGTATGCAATTCGCGAAGCAGTAGTATACCATCAGGCTCTATTTC 680

QY      741 CCCCCANCTTGGGAAAAACGGGATNGTTCCTCCCCCTT---AAAAACAACCTTCCCCCN 797
      |||||
Db      679 ACCCAGACATAGGAAAAACAGGATTGTCACACCCATTAAGAAACAATCCTATCCACG 620

QY      798 CTTTGCCCGCAGGNNNTTTCCTCCGCTTAATCCGAACATTAATAAG 843
      |||||
Db      619 CTATTTGCCAATGTGTAGTTTCACTTAATTTCTGACATAATAAATG 574

RESULT 15
HUMSCP2B/c
LOCUS
DEFINITION Human sterol carrier protein 2 mRNA, complete cds.
ACCESSION M75884
VERSION M75884.1 GI:432976
KEYWORDS sterol carrier protein-2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1500)
AUTHORS He,Z., Yamamoto,R., Furch,E.E., Schantz,L.J., Naylor,S.L.,
George,H., Billheimer,J.T., and Straus,J.F., III.
CDNAs encoding members of a family of proteins related to human
sterol carrier protein 2 and assignment of the gene to human
chromosome 1 p21---pter
JOURNAL DNA Cell Biol. 10 (8), 559-569 (1991)
MEDLINE 92029618
PUBMED 1718316
REFERENCE 2 (bases 1 to 1500)
AUTHORS Vega,J., Hellsten,E., Branoski,B.L., Emanuel,B.S., Billheimer,J.T.,
Mead,S., Cowell,J.K., Straus,J.F., III, and Peltonen,L.
Assignment of sterol carrier protein X/sterol carrier protein 2 to
1p32 and exclusion as the causative gene for infantile neuronal
ceoid lipidufionosis
JOURNAL Unpublished
COMMENT On Dec 6, 1993 this sequence version replaced gi:337996.
FEATURES
  source
    1.1500
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /map="1p32"
      /issue_type="liver"
      /clone_lib="Clontech HL-1001b;HL-1115b"
      1..1500
      /gene="SCP-2"
      82..951
      /gene="SCP-2"
      /codon_start=1
      /product="sterol carrier protein-2"
      /protein_id="AA03558.1"
      /db_xref="GI:432977"
      /translation="MTDLPSFEEKSIKRVGFDMSCKEARKCYEKSGLTPNDIDVI
      ELHDCFSTNELTYEALGLCPGOGATLVDRDNTYGGKVINPSSGLISKHPLGAT
      GLAOCALCWLORGEAGKROVPGAKVALQHNLDIGAVVVTLYKMGPFAASSFRTHQ
      IEAVPTSSASDGFKANLVFKEIEKLDEEGQFVKKIGIFAFKVDGPGKEATWV
      DVKNGKSVLPNSDKADCTITMADSDFLAINTGKNPQSAFPGKLIKITGNMGLAMK
      LQNLQIPGNAKL"
      1218..1223
      /gene="SCP-2"
      /note="no polyA or polyA signal was found at the end of
      polyA_signal
```

this cDNA clone, indicating that the mRNA sequence may be

partial at the 3' end"

BASE COUNT 482 a 254 c 342 g 422 t
ORIGIN

Query Match 31.2%; Score 262.8; DB 9; Length 1500;
Best Local Similarity 70.9%; Pred. No. 5.2e-44;
Matches 348; Conservative 0; Mismatches 137; Indels 6; Gaps 3;

```
QY 359 TTGGAATGGGAAAAAATNCNGTCCNNATACCCNNGAANNCAAANTTTTAAATTTT 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1500 TTTGATTTGGGAAGAAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTT 1441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 TAACCCCTTANTTTTAAANCTATNGAAANTNGATTANNACTTGAATTGC-CAACCC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1440 TAATCTCACTAATTTTAAAGACTATGAGAAATGATTAATGACATGAAGTCACACAC 1381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 TANTTNCNGGCCACCCNGTGGGCTNGTNTTCTTACTTANTCCCCCAAGAAANCCCTT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 TAATTAAGGCCAGCTGTGTCATGTGTTCTTACTTACTTCTCCCAAGGAAACTCTT 1321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 AANCGAANCTCCNCCAAATAACCTTAANTATCCTTGTTAACCAANCAAAACCTTTT 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 AAACGTGAATCTTCAGCAGAAATATCTTAATATATACTTGTGAAGCAAAACAAAGCTTT 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 598 TNGTTACNTANTCCTTGGATTTAAGGGTCCCAATTNATCGAACCANTTTTCC 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1260 TTGTTTACATAGTCTTGTGATTTTACTGTTCTTAATTTTATCTGAAGCAATTTTAC 1201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 658 CCCNAACCATANTTACCATTTTACCTTGTGAAGCNCAGTNGTTCANTNCCGCAAANC 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1200 CCCAGACCATAATTACCATATTTAATTGTAATGCACAGTTGTATGCAATTCCGCAAAGC 1141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 718 AGTANTNTTC--CCNGGCTTTCCTCCCGGAGNCCTTGGAAGAAACGGGATNGTCCCCC 775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1140 AGTAGTATACCATCAGGCTCTATTCACCCAGACATAGAGAAACAGGATTGGTCACACC 1081
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 776 CTT--AAAAACAACCTTCCCCCCTTTGGCCCCAGGNNNTTTCCTTAATCCGA 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1080 CATTAAGAAACAAATCTATCCACGCTATTGCAATGTGTAGTTTCAAGTCTAAATCTG 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 833 ACAATAAAAAG 843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1020 ACAATAAATG 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: November 27, 2003, 10:02:41
Job time : 3416.44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:55:53 ; Search time 2272.04 Seconds
(without alignments)
9017.749 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 843
Sequence: 1 ggccaaaanaaaatcatttna.....taaatccgaacataaaaag 843

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vtr: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rtd: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483.4	57.3	786	9	AM052045 wx25e05.x
2	483	57.3	953	9	AL578584 AL578584
3	479.4	56.9	901	13	BX392565 BX392565
4	477.6	56.7	821	9	AI640146 wa299f11.x

5	472.2	56.0	756	12	BM985376	BM985376 UI-CF-EC1
6	472.2	56.0	761	12	BM997078	BM997078 UI-H-ED0-
7	468.2	55.5	769	12	BQ014192	BQ014192 UI-H-ED1-
8	466.2	55.3	905	13	BX414370	BX414370 BX414370
9	466	55.3	770	9	AI826287	AI826287 wk33f07.x
10	463.4	55.0	793	9	AI565988	AI565988 cn52a05.x
11	462.2	54.8	762	13	BU627147	BU627147 UI-H-FG0-
12	460	54.6	824	10	BE738457	BE738457 601572489
13	459	54.4	759	13	BU619112	BU619112 UI-H-FH1-
14	458.2	54.4	822	9	AV7229461	AV7229461 AV7229461
15	455.6	54.0	738	12	BQ045161	BQ045161 UI-CF-EN1
16	455.4	54.0	750	12	BQ009270	BQ009270 UI-H-ED1-
17	455.2	54.0	734	14	CA424156	CA424156 UI-H-FE1-
18	454.8	54.0	840	9	AA659242	AA659242 nu12h08.s
19	453.6	53.8	837	9	AI566109	AI566109 cn53e08.x
20	452.8	53.7	784	12	BI861076	BI861076 603390807
21	450.2	53.4	751	12	BQ008197	BQ008197 UI-H-ED1-
22	449	53.3	734	13	BU933572	BU933572 AGENCOURT
23	448.4	53.2	721	12	BM968746	BM968746 UI-CF-DU1
24	447.4	53.1	1032	9	AL551280	AL551280 AL551280
25	442.6	52.5	712	13	BU619082	BU619082 UI-H-FH1-
26	441.8	52.4	876	9	AI831751	AI831751 wj40a10.x
27	440.8	52.3	820	10	BF966918	BF966918 602286426
28	438.2	52.0	699	10	BE739622	BE739622 601556314
29	432.8	51.3	714	9	AI675901	AI675901 wc06h04.x
30	427	50.7	672	12	BQ016778	BQ016778 UI-H-D10-
31	426.4	50.6	1201	9	AL577351	AL577351 AL577351
32	424.2	50.3	1018	10	BE379783	BE379783 601159529
33	421.2	50.0	749	9	AI765940	AI765940 wh69d12.x
34	414.4	49.2	728	9	AI831053	AI831053 wj62c05.x
35	410	48.6	1201	9	AL519776	AL519776 AL519776
36	409.8	48.6	648	12	BM311898	BM311898 ig63e08.y
37	409.2	48.5	671	9	AI683310	AI683310 tw50e02.x
38	408.4	48.4	673	9	AW512844	AW512844 xm04g12.x
39	407	48.3	744	10	BG569206	BG569206 602588480
40	406.8	48.3	1201	13	BX444130	BX444130 BX444130
41	403.4	47.9	942	10	BE614293	BE614293 601504344
42	403	47.8	637	13	BQ549593	BQ549593 ik88g10.x
43	402.6	47.8	763	9	AI889069	AI889069 td17a12.x
44	401.6	47.6	641	9	AW173415	AW173415 xj06d03.x
45	400	47.4	616	9	AW611897	AW611897 hg90b03.x

ALIGNMENTS

RESULT 1
AM052045 786 bp mRNA linear EST 20-SEP-1999
LOCUS wx25e05.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2544704 3'
DEFINITION similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION AM052045
VERSION AM052045.1 GI:5914404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.lnl.gov/bhrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 458.

FEATURES

source

1.786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2544704"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: kidney; Vector: pT7F3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss-circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 290 a 116 c 109 g 270 t 1 others
ORIGIN

Query Match 57.3%; Score 483.4; DB 9; Length 786;
Best Local Similarity 72.1%; Pred. No. 2.9e-79;
Matches 567; Conservative 0; Mismatches 216; Indels 3; Gaps 2;

QY 1 GGCCAAAAAANTTATTNNAATTCCTATTANCNTCTCCNCAANCAATTATTNACC 60
Db 1 GGCCAGAAAAAGTATTATTCTATTAAACATTCTGCGAAAAGCATTATTATCC 60
QY 61 TATNNCNCNGANTTTNANAANTACCTTNTNTTAAAAAACCCTNGAAAAAAATAA 120
Db 61 TATATCTCACTGAATTTTAAAGAAATAACATTAGTATTAGAAAAAAGATAA 120
QY 121 TNGCAATANTTAACCTTNCCTGAAANGAAATTTNTACCAANGAANGAANCNTNT 180
Db 121 ATGCAGATAATTAACCTTACATGAAAAAGAAATATATAACAAGACTGAGAACGTTAT 180
QY 181 AATTNGAANTNAAATTATANTNGAANGCGCNCNGAAGCAACACTTATGTCCAATT 240
Db 181 AATTGAAATGAGATTATATTTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAAAT 240
QY 241 ATCTTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTT 300
Db 241 ATTCTTAATGATGTGTTTATGACTAATACACTGATTTTCAATPAAGAAACCCATGTT 300
QY 301 AAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCTTNAAT 360
Db 301 AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAT 360
QY 361 TGGATTGGGAAAAAATNCNGTCCNNATACCCNNGAANGCAANTTTTAAATTTTA 420
Db 361 TTGATTTGGGAAAGAAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTA 420
QY 421 ACCCCCTANTTTTAAANCTAINGAAAAANTNGATTANNGACTGATTC-CAACCTA 479
Db 421 ATCTCACTAATTTTAAAGACTATGAGAAATGATTATGACATGAAGTGACACAACACTA 480
QY 480 NTTNNGGCCACCGTGGCCTNGTNTTCTTACTTANTCCCCCAAGGAAANNCCTTAA 539
Db 481 ATTACTGGCCAGCTGTGGCATGTGTTCTTACTTAGTTCTCCCAAGGAAACTTTAA 540
QY 540 NCGAANCTCCNCCAAATAACCTTAANTATCCTTGTAACCAANCAAAACCTTTTN 599
Db 541 ACTGAATCTTCAGCAGAAATATCTTAATATATCTTTGTAAGCAAAACAAAGCTTTT 600
QY 600 GTTACNTANTCCTTGGATTTAACGGGTCCCAATTTNATCCNGAACCAANTTTCCCC 659
Db 601 GTTACATAGTCTTTGATTTTACTGTTCCTAATTATATTCTGAACCTCAATTTTACC 660
QY 660 TNAACCATANTTACCATTCTTGTAAGGCNAGTNGTTGCANTNCCGCAANAG 719

Db 661 CAGACCATATTAATTAACCTTTGTAATGACAGTGTATGCAATTNCGAAAGCAG 720
QY 720 TANTNTTCC--CCNGCNCCTTTCCCGGANCCTTGGGAAAAAACGGATNGTCCCCCT 777
Db 721 TAGTATACCATCAAGCTCTATTCACCCAGAGCATAGGAAAGACAGATTGTTCAACCCA 780
QY 778 TAAAAA 783
Db 781 TAAGAA 786

RESULT 2
AL578584
LOCUS
DEFINITION AL578584 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK001YL01 3-PRIME, mRNA sequence.
ACCESSION AL578584
VERSION AL578584.2 GI:31316785
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS 1 (bases 1 to 953)
TITLE Li W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT On Feb 16, 2001 this sequence version replaced gi:12942799.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 268.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK001CF01NP1&cluster=268.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK001CF01NP1.

FEATURES

source

1.953
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YL01"
/cell_line="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 336 a 145 c 136 g 325 t 11 others
ORIGIN

Query Match 57.3%; Score 483; DB 9; Length 953;
Best Local Similarity 70.4%; Pred. No. 3.4e-79;
Matches 584; Conservative 7; Mismatches 232; Indels 6; Gaps 3;

QY 21 ATTTCCTATTANCTCTCCNCAANCAATTATTNACCCTATNNCNCNGANTTNAN 80
Db 3 AKKKGHNTTAAACATCTCTCAAGACATATTATCTCTATATCTCACTGAATTTAA 62
QY 81 AAANTACCTTNTNTTAAAAAACCTNGGAAAAAAATAATNGCAATANTTAACCTTNC 140
Db 63 GAAATACATTAGTATTGAAAAAACTAGGAAAAAAGATTAATGACATTAATTAACCTTAC 122
QY 141 TTGAAAAANGAAATTTNTTACCAANGACNGAANCNTNTTAATTNGAANTNAAATTATAN 200
Db 123 ATGAAAAAGGAAATTATTAACAAGAGACTGGAACGTTATAAATTGAAATGAGATTATAA 182

QY	201	TTNGAANCGGCNNCNGAABCAANCTTNATGGTCCAAATTACTCTNAANGAGGNNTTN	260
DB	183	TTGAAAACCTGCATCTGAAAGCAAACTTATGTGTTCAATTATCTTAATGATGGTGT	242
QY	261	ANNACTAATNCCNGATTTTCCAATANGAANCCNNNTTAAANTNTTTNATTTTAA	320
DB	243	ATGACTAATACACTGATTTTCAATAAGAAACCAGTGTAAATAATTTTATTTTAA	302
QY	321	AATAACCCNGTNTCCAACCCNGATCANATTCTTTNATTTGGATTGGGAAAAAATNC	380
DB	303	AATAAGCCGTGTTTCAAGCTGATCATATTTCTTTATTTGATTGGAGAAGAAATAC	362
QY	381	NGTTCNMATACCNNGAANNGCAANTTTTAAATTTTAAACCCCTANTTTTAAAN	440
DB	363	TGTTTCTGATAGCATGAATGCAAAATTTTGAATTTTAACTCACTAATTTTAAGAC	422
QY	441	TATNGAAAANTNGATTANNGACTGGAATTGC-CAACCTTANTTNCNGCCACCNGTGGC	499
DB	423	TATTGAGAAATGATTAAATGACATGAAGTGACACAACATAATTACTGGCCAGCTGTGGC	482
QY	500	NTNGNTTCCCTTACTTANTCCCCCAAGAAANNCCTTAANGAANCTCCNCCAAAATA	559
DB	483	ATTGTTTTTCTTACTTAGTTCTCCCAAGAAAACCTTAACTGAATCTTCAGAGAATA	542
QY	560	ACCCTAANTATCCTTGGTAACCAANCAAAAACCTTTTNGTTTACNTANTCCTTGGAT	619
DB	543	ATCCTTAATATACTTGTGTAAGCMAAAACAAAAGCTTTTGTTTTACATAGTCTTGGAT	602
QY	620	TTAACGGGTCCTCAATTNATCCNGAACCCANTTTTCCCCCNAAACATANTTACCATTT	679
DB	603	TTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACCAGANCAATAATTACCATATT	662
QY	680	ACCTTGTAAGGCNCAGTNGTTGCANTNCCGCAANCAAGTANTNTTC--CCNGGCNCT	737
DB	663	AACTTTGTAATGACAGTGTATGCAATTYCGCAAAGCAGTGTATACCATCAGGCTCTA	722
QY	738	TTCCCCCGANCCTTGGGAAAAACGGGATNGTTCCTCCCTT--AAAAACAACCTTCCC	794
DB	723	TTCAACCAAGACATAGGAAAAACAGGATTGTCACACCCATTAAAGAAACAATCTATCCC	782
QY	795	CCNCTTTGGCCCAAGNNTNTTCCCGTCTAATCCGAACAATAAAAG	843
DB	783	ACGCTATTGGCCAAATGTGTAGTTTCACTTAATTTCTGACAAATAAATG	831

RESULT 3
BX392565
LOCUS
DEFINITION
BX392565 901 bp mRNA linear EST 13-MAY-2003
BX392565 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1022YF01 3-PRIME, mRNA sequence.
ACCESSION
BX392565
VERSION
BX392565.1 GI:30607982
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 901)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
Contact: Genoscope
AUTHORS
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
TITLE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
COMMENT
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 268.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAK054BD07NM1&cluster=268.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BAK054BD07NM1.

FEATURES		Location/Qualifiers	
Source		1. .901	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0D1022YF01"	
		/issue_type="PLACENTA COT 25-NORMALIZED"	
		/clone_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"	
		/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
BASE COUNT	326 a 139 c 123 g 309 t 4 others		
ORIGIN			
Query Match	56.9%;	Score 479.4;	DB 13; Length 901;
Best Local Similarity	70.2%;	Pred. No. 1.6e-78;	
Matches 592;	Conservative 0;	Mismatches 245;	Indels 6; Gaps 3;
Qy	5	AAAAAAANTTATTNAATTTCTCTATTAAACNTCTCCNCAAAACATTATTNACCTATN	64
Db	7	AAAAAGCTATTTTAAATTTCTATTAAACATCTCTTAAAGCATTATTATCTTATA	66
Qy	65	NCNCNGANTTTNANAANTACCTTNTNTTTAAAAAACCCTNGGAAAAAATAATNGC	124
Db	67	TCTCACTGAATTTTAAAGAAATAACATTAGTATAGAAAAAAGATAAATGC	126
Qy	125	AAATANTTAACCTTCTTGAAANGGAATTTNTACCAANGGACNGAANCNTTNTAAT	184
Db	127	AGATAATTAACTTACATGAAAAAGGAAAAATTATTAACAAGGACGTAGAACGTTATAAAT	186
Qy	185	NGAANTNAAATTATANTNGAAANCGGCNNCNGAAACCAANTTNATGTGTCATTAATCC	244
Db	187	TGAAATGAGATTATATTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATTATTC	246
Qy	245	TNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGGAANCCNNNTTAAAA	304
Db	247	TTAATGATGTGTTTATGACTAATACACTGATTTTCAAGAAGAAACCATGTATAAA	306
Qy	305	NTNTTTTNAATTTTAAAAATAACCCNGTNTCCAACCCNGATCANATTCCTTNTATTGA	364
Db	307	ATATTTTATTATTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTTATTGTA	366
Qy	365	TTGGGAAAAAATNCNGTTCNNATACCCNNGAANNGCAANTTTTAAATTTTAAACC	424
Db	367	TTTGGAGAAAAATACGTCTTCTGATAGCATGAAATGCAAAATTTTACTTTTAAATCT	426
Qy	425	CCCTANTTTTAAANCTATNGAAANTNGATTANNGACTGAATTGC-CAACCTANTTN	483
Db	427	CACATAATTTTAAAGACTATGAGAAATGTATTATGACATGAAGTGACACACACTAATTA	486
Qy	484	CNGGCCACCGTGGGCTNGTNTTCTTACTTANTTCCCCCAAGGAAANNCCTTAANCNG	543
Db	487	CTGGCCAGCTGTGGCAATGTGTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAATTG	546
Qy	544	AANCTCNCNCAAAATTAACCTTAANTATCTTGGTAAACCAANCAAAACCTTTTNGTTT	603
Db	547	AATCTTCAGCAGAAATATCTTAAATATACTTTGTGAACAAAAAGCTTTTGTGTTT	606
Qy	604	ACNTANTCCTTGGATTTAACGGGTCGCCCAATTNTATCCNGAACCCANTTTTCCCCNAA	663
Db	607	ACATAGTCTTTGGANTTTACTGTTCTTAATTTTATCTGAAACTCAATTTTACCCGAGA	666
Qy	664	CCATANTTACCATTTTAACTTGTGTAAGCNCAGTNGTTTGCANTNCCGCAAAACAGTANT	723
Db	667	CCATANTTACCATATTAACCTTGTATATGACACAGTTGTATGCAATTCGCAANAGCAGTAGT	726
Qy	724	NTTCCC--CNGGCNCTTTCCCCCGANCCCTTGGGAAAAACGGGATNGTCCCCCCCC--TT	778
Db	727	ATACCCATCAGCTCTATTCACCCAGAGCATAGGAAAAACAGGATTGGTCACACCCCAATTA	786
Qy	779	AAAAACAACCTTCCCCCNCTTTGGCCACAGNNNTTTCGGCTTAATCCGAACAATA	838

Db 787 GAAACAATCTATCCACGCTTTGGCCAATGTGAAGTTTCAGTCTAAATTCTGACATAA 846

QY 839 AAA 841

Db 847 AAA 849

RESULT 4

LOCUS A1640146 821 bp mRNA linear EST 16-DEC-1999

DEFINITION wa29f11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:229533 3' similar to gb:S52450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION A1640146

VERSION A1640146.1 GI:4703255

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 821)

TITLE NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

COMMENT Unpublished

CONTACT: Robert Strausberg, Ph.D.

EMAIL: cgapbs-r@mail.nih.gov

TISSUE Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI_CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: www-bio.1nl.gov/bbrp/image/image.html

Insert Length: 565 Std Error: 0.00

Seq primer: -40UP from G1bco

High quality sequence stop: 471.

FEATURES

source

location/Qualifiers

1..821

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:229533"

/lab_host="DH10B"

/clone_lib="NCI CGAP_Kid11"

/note="Organ: kidney; Vector: pT713D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 296 a 128 c 114 g 278 t 5 others

ORIGIN

Query Match 56.7%; Score 477.6; DB 9; length 821;

Best Local Similarity 70.9%; Pred. No. 3.4e-78;

Matches 581; Conservative 0; Mismatches 235; Indels 4; Gaps 3;

QY 2 GCCAATAAANTTATTNAATTCCTATTAAACNTCCGCAACATTTATTTNACCT 61

Db 1 GCCAGAAAAAGTATTATTCTATTAACAATCTTCACAAGCATTTATTTATCCT 60

QY 62 ATNNCNCNGANTTTNANAANTACCTTTNNTNTTAAAAAACCTNGGAAAAAATAAT 121

Db 61 ATATCTCACTGAATTTTAAAGAAATACATTAGTATTAGAAAAAGATAAAA 120

QY 122 NGCAATANTTAACCTTCTTGAAAAANGAAATTNTTACCAANGAAGAAANCNTNTA 181

Db 121 TGCAGATAATTAAACTTACATGAAAAAGAAAAATTATACAAAAGACTGAGACGTATA 180

QY 182 ATTNGAANTNAAATTATANTNGAANCGGCGNNGAAGAACCAANCTNATGGTCCAATTA 241

Db 181 AATTGAATGAGATTATATTGAAAGTCACTGGAAGCAAACTTATGTGTAATTA 240

QY 242 TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCATANGAANCCNNNTTA 301

Db 241 TTCTTAATGATGGTGTTTATGACTAATACACTGATTTTCAATAAGAAACCATGTTA 300

QY 302 AAANTNTTTNATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTTNATTT 361

Db 301 AAAATATTTTATTATTTAAAAATAAGCCTGTGTTCAGAGCTCTGATCATATTTCTTTATTT 360

QY 362 GGATTGGGAAAAAATNCGNTTCCNNATACCNNGAANNGCAAANTTTTAAATTTTAA 421

Db 361 TGATTTGGAGAAAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTAA 420

QY 422 CCCCCCTANTTTTAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCCCTAN 480

Db 421 TCTCACTAATTTTAAGACTATGAGAAATGATTAATGACATGAAGTGACACACTAA 480

QY 481 TTNCNGGCCACCGTGGCGNTNGTNTCTCTTACTTANTTCCCCCAAGAAANNCTTAAN 540

Db 481 TTACTGGCCAGCTGTGTGGCATTTGTGTTCTTACTTAGTCTCCCAAGAAACTCTTAA 540

QY 541 CNGAANTCCNCCAAATAACCCTTAANTATCTTGTGTAACCAANCAAAACCTTTTNG 600

Db 541 CTGAATCTTCAGCAGAAATAATCCTTAATAATACTTTGTAAGCAAAACANAGCTTTNTG 600

QY 601 TTTACNTA-NTCCTTGGATTTAACGGGTCCCAATTNATCCNGAACCAANTTTTCCC 659

Db 601 TTTACATAGNTCTTTGGATTTTACTGTGCTTAATTTTATTTCTGAACCTCAAAATTTACC 660

QY 660 CNAACCATANTTACCAATTTTACCTTGTAAGCAGCAGTNGTTGCANTNCCGCAANACAG 719

Db 661 CAGACCATAATTACCATNATAACNTTTGTATGACAGAGTGTATGCAATTCGCGCAAGCAG 720

QY 720 TANTNTTCC--CCNGGNCCTTTCCCCCGANCTTGGGAAAAACGGATNGTCCCCCT 777

Db 721 TAGTATACCATCAGGCTCTATTTCACCCAGACGATAAGAAACAGATGGGCCACACCAT 780

QY 778 TAAAAACAACCTTCCCCCNCCTTTGGGCCAGANNNTNTT 817

Db 781 TAGAAAACAATCTATCCACGCTATTTGCCCCAGTGTA 820

RESULT 5

LOCUS BM985376 756 bp mRNA linear EST 20-FEB-2003

DEFINITION UI-CF-Ec1-acg-p-23-0-UI.s1 UI-CF-Ec1 Homo sapiens cDNA clone

UI-CF-Ec1-acg-p-23-0-UI 3', mRNA sequence.

BM985376

BM985376.1 GI:19611803

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 756)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 316-342, >AT_rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source

Location/Qualifiers
1. .756
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acg-p-23-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGTGCTTAC.
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGTGCTTAC"

BASE COUNT 268 a 107 c 97 g 283 t 1 others
ORIGIN

Query Match 56.0%; Score 472.2; DB 12; Length 756;
Best Local Similarity 73.2%; Pred. No. 3.4e-77;
Matches 533; Conservative 0; Mismatches 194; Indels 1; Gaps 1;

QY 2 GCCAAAAAANTATTATTTCTTATTAACNTCCCTCCNCAANCAATTATTTACCT 61
Db 19 GCCAGAAAAAGTATTATTTCTTATTAACATTTCTTCTCAAGCATTTATTCCT 78
QY 62 ATNNCNCNCNGANTTTNANAANTACCTTTNTNTTAAAAAACCCTNGAAAAAATAAT 121
Db 79 ATATCTCACTGAATTTAAGAAATACATTAGTATTAGAAAAAAGTAA 138
QY 122 NGCAATANTTAACCTTNCCTGAAAAANGAAATTTTACCAANGAGCNGAAANCNTNTTA 181
Db 139 TGCAGATAATTAACTTACATGAAAAAGAAATTTATACAAAGAGCTGAGAACGTTATA 198
QY 182 ATNGAANTNAAATTATANTNGAAANGCGCNCNCAACCAANCTTNATGTCATTA 241
Db 199 AATTGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTTATGTTCAATTA 258
QY 242 TCCTNAANGAGGNNTTNANNACTAATNCCCGATTTTCCAATANGAANCCNNTTA 301
Db 259 TTCTTAATGATGTGTTTATGACTAATACACTGATTTTCAATAGAAACCACTGTTA 318
QY 302 AAANTNTTTTATTTAAATAAACCNGTNTCCAACCCCGATCANATTCCTTNATTT 361
Db 319 AAAATATTTTATTTAAATAAAGCCGTGTGTTCAAGCTTGATCATATTTCTTTATTT 378
QY 362 GGATTGGGAAAAAATNCNGTCCNNATACCNNGAANNCAANTTTTAAATTTTAA 421

Db 379 TGATTGGGAAGAAATACTGTTTCTGATAGCATGAATGCAAAATTTTATGATTTTAA 438
QY 422 CCCCCCTANTTTTAAANCTATNGAAAAANTNGATTANNGACTGAATTGC-CAACCTTAN 480
Db 439 TCTCACTAATTTTAAAGAACTATTGAGAAATGATTAATGACATGAAGTGACACACTAA 498
QY 481 TTNCNGCCACNGTGGCGCNTNGTNTTCTTACTTANTCCCCCAAGAAANNCTTAA 540
Db 499 TTTACTGGCCAGCTGTGGCATTTGTGTTCTTACTTAGTTCTCCCAAGAAACTTTAA 558
QY 541 CNGAANTCCNCCAAAAATAACCCTTAANTATCTTGTGTAACCAANCAAAACCTTTTNG 600
Db 559 CTGAATCTTCAGCAGATAATATCTTAAATATACTTTGTAAGCAAAACAAAGCTTTTGG 618
QY 601 TTTACNTANTCTTGGGATTTTAACGGGTCCCAATTNATCCNGAACCCANTTTTCCCC 660
Db 619 TTTACATAGTCTTTGGATTTTACTGTTCCCTTAATTTTATTCGAAACTCAATTTTACCC 678
QY 661 NAACCATANTTACCAATTTTACCTTGTAAGCNGCAGTNGTTGCANTNCCGCAANCACT 720
Db 679 AGACCAATATATACATATTAACCTTGTGAATGACAGAGTGTATGCAATTCGCAANAGCAGT 738
QY 721 ANTNTTCC 728
Db 739 AGTATACC 746

RESULT 6
BM997078 761 bp mRNA linear EST 26-MAR-2002
LOCUS
DEFINITION
UI-H-ED0-axo-f-03-0-UI.s1 NCI CGAP_ED0 Homo sapiens cDNA clone
IMAGE:5831426 3', mRNA sequence.
ACCESSION
BM997078
VERSION
BM997078.1 GI:19721979
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 761)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaphs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
The following repetitive elements were found in this cDNA
sequence: 316-342, >AT_rich#low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
Source
Location/Qualifiers
1. .761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5831426"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_1lb="NCI_CGAP_ED0"
/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac
(Pharmacia) with a modified polylinker; Site_1: EcoR I;
Site_2: Not I; NCI CGAP_ED0 is a cDNA library containing
the following tissue(s): Chondrosarcoma cell line C55. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer

containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GCTCAAGGCT.

TAG LIB=UI-H-ED0
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"

BASE COUNT 269 a 108 c 98 g 285 t 1 others
ORIGIN

Query Match 56.0%; Score 472.2; DB 12; Length 761;
Best Local Similarity 73.2%; Pred. No. 3.4e-77;
Matches 533; Conservative 0; Mismatches 194; Indels 1; Gaps 1;

QY 2 GCCAATAAANTTATTTTAAATTCCTTAATTAAACNTCCCTCCNCAANCAATTATTNACCT 61
Db 19 GCCAGAAAAGTATTATTAAATTTCTATTAAACATCTCTCAAGCATTTATTATCTCT 78
QY 62 ATNNCNCNGANTTTNANAANTACCTTNNNTTAAAAAACTNGAAAAAAATAAT 121
Db 79 ATATCTCACTGAATTTTAAAGAAATACATTAGTATTAGAAAACTAGAAAAAGATATA 138
QY 122 NGCAATANTTTAACCTTNCCTGAAANGAAATTTNTACCAANGACNGAANCNTNTTA 181
Db 139 TGCAGATAATTAACTTACATGAAAAAGAAATTTATACAAAGAGACTGAGACGTATA 198
QY 182 ATTNGAANTNAATTTATANTTNGAANGCGCNCNGAAGCAACCAANTTNATGTCATTA 241
Db 199 AATTGAAATGAGATTATATATTGAAAACTGCATCTGAAGCAAACTTATGTCTCAATTA 258
QY 242 TCCTNANGAGGNNTTNANNACTAATNCCNGATTTTCCAAATANGAANCNNNTTA 301
Db 259 TTCTTAATGATGTGTTTATGACTAATACTGATTTTCAATAAGAAACCCATGTTA 318
QY 302 AAANTNTTTNATTTTAAAAATACCCNGTNTCCAACCCCAATCANATTCCTTNNATTT 361
Db 319 AAAATATTTTATTAAATAAAGCTGTGTTCAAGCTTGATCATATTTCCTTAAATTT 378
QY 362 GGATGGGGAATAATNCCNGTTCNNATACCNNGAANGCAANTTTTAAATTTTAA 421
Db 379 TGATTTGGGAAAGAAATACGTCTTGATAGCATGAATGCAAAATTTTAGATTTTAA 438
QY 422 CCCCCCTANTTTTAAANCTATNGAANAANTNGATTANNGACTGAATTGC-CAACCTTA 480
Db 439 TCTCACTAATTTTAAAGAACTATGAGAAATGATTATGACATGAGTGCACACACTAA 498
QY 481 TTNCNGGCCACCCNGTGGCCTGTTCTTACTTANTCCCCCAAGGAANNCTTAAN 540
Db 499 TTAAGGCCAGCTGTGGCATGTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAA 558
QY 541 CNGAANTCCNCCAAAATAACCTTAANTATCTTGGTACCAAAACAACCTTTTNG 600
Db 559 CTGAATCTTCAGCAGAAATATCTTAATATACTTGTAAAGCAAAACAAGCTTTTGTG 618
QY 601 TTTACNTANTCCTGGGATTAAAGGGTCCCAATTTNATCNGAACCANTTTTCCCC 660
Db 619 TTTACATAGTCTTTGGATTTAAGTGTCTTCTAATTTTATTTGAAAACCTCAATTTTACC 678
QY 661 NAACCATANTTACATTTTACCTTGGTAAAGCNCAGTNGTTGCANTNCCGCAANCACT 720
Db 679 AGACCATATATTAACATATTAACCTTGGTATGACACAGTGTATGCAATTCGCGANAGCAGT 738
QY 721 ANTNTTCC 728
Db 739 AGTATACC 746

RESULT 7
BQ014192
LOCUS BQ014192 769 bp mRNA linear EST 26-MAR-2002

DEFINITION UI-H-ED1-axs-g-24-0-UI.s1 NCI CGAP_ED1 Homo sapiens cDNA clone
IMAGE:5833007 3', mRNA, sequence.
ACCESSION BQ014192
VERSION BQ014192.1 GI:19739093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
1 (bases 1 to 769)
AUTHORS NCI-CGAP
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LIN at: <http://image.lnl.gov>
The following repetitive elements were found in this cDNA
sequence: 317-343, >AT_rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1..769
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5833007"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_ED1"
/note="Organ: Left Pubic Bone; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP_ED1 is a normalized cDNA library
containing the following tissue(s): Chondrosarcoma cell
line C85. The library was constructed according to Bonaldo
, Lennon and Soares, Genome Research, 6:791-806, 1996.
First strand cDNA synthesis was primed with an oligo-dT
primer containing a Not I site. Double stranded cDNA was
ligated to an Ecor I adaptor, digested with Not I, and
cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tag for this library is GCTCAAGGCT.
TAG LIB=UI-H-ED1
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"

BASE COUNT 272 a 110 c 100 g 285 t 2 others
ORIGIN

Query Match 55.5%; Score 468.2; DB 12; Length 769;
Best Local Similarity 72.7%; Pred. No. 1.8e-76;
Matches 543; Conservative 0; Mismatches 202; Indels 2; Gaps 2;

QY 1 GCCAATAAANTTATTTTAAATTCCTATTTAAACNTCCCTCCNCAANCAATTATTNACCC 60
Db 19 GCCAGAAAAGTATTATTAAATTTCTATTAACATTTCTCTCAAGCATTTATTATCC 78
QY 61 TATNNCNCNGANTTTNANAANTACCTTNNNTTAAAAAACTNGAAAAAAATAA 120
Db 79 TATATCTCACTGAATTTTAAAGAAATACATTAGTATTAGAAAAAACTAGAAAAAAGATTA 138
QY 121 TNGCAATANTTTAACCTTNCCTGAAAAANGAAATTTNTACCAANGACNGAANCNTTNT 180
Db 139 ATGCAGATAATTAACTTACATGAAAAAGAAATTTATTAACAAAGAGACTGAGACGTAT 198
QY 181 AATTNGAANTNAATTTATANTTNGAANGCGCNCNGAAGCAACCAANTTNATGTCCAATT 240

Db 199 AAATGAAATGAGATTATATTTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATT 258
Qy 241 ATCTNNAANGAGGNNNTTANNACTAATNCCNGATTTCCATANGAANCCNNNTT 300
Db 259 ATTCTTAATGATGTTGTTTATGACTAATACACTGATTTTCAATGAAGAAACCATGTT 318
Qy 301 AAAAATNTTTTATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAAT 360
Db 319 AAAAATATTTTATTTTAAAAATAGCCTGTGTTCAAGCTGTGATCATATTTCTTTAAT 378
Qy 361 TGGATTGGGAAAAAATNCNGTTCNNATACCCNNGAANNGCAAANTTTTAAATTTTAA 420
Db 379 TTGATTTGGGAAGAAATACGTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 438
Qy 421 ACCCCCTANTTTTAAANCTATNGAAANTNGATTANNGACTTGAATTGC-CAACCTA 479
Db 439 ATCTCACTAATTTTAAAGAACTATGAGAAATTGATTATGACATGAGTGCACAACTA 498
Qy 480 NTNNGNGCCACCGNGTGGCNGTNTTCTTACTTANTCCCCCAAGGAAANNCCTTAA 539
Db 499 ATTAAGTGGCCAGCTGTGGCATTTGTTCTTACTTACTTCTCCCAAGGAAACTCTTAA 558
Qy 540 NCNGAANTCCNCCAAATAACCTTAANTATCCTTGTAACCAANCAAAACCTTTTN 599
Db 559 ACTGAATCTTCAGCAGAAATATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTT 618
Qy 600 GTTACNTANTCCTTGGAATTAAGGGTCCCCAATTNATCCNGAACCANTTTTCCC 659
Db 619 GTTACATAGTCTTGTGATTTACTGTTCTTAATTTATTCGAAACTCAATTTTACC 678
Qy 660 CNAACCATANTTACCATTTTACCTTGTAAGGCNCAAGTNGTTGCANTNCCGCAANCA 719
Db 679 CAGACCATATTAATCATATTAATCTTGTNTATGACAGTTGTATGCAATT-CGCANAGCAG 737
Qy 720 TANTNTTCCCCNGCNCCTTCCCCCGA 746
Db 738 TAGTATACATCAGGCTTATTCACCCA 764

RESULT 8
BX414370 905 bp mRNA linear EST 13-MAY-2003
LOCUS BX414370 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM007YJ07 3-PRIME, mRNA sequence.
ACCESSION BX414370
VERSION BX414370.1 GI:30637003
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 905)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 268.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAK077DD07NM1&cluster=268.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0BAK077DD07NM1.
FEATURES
Source location/Qualifiers
1..905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="CS0DM007YJ07"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: Liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 333 a 136 c 123 g 312 t 1 others
ORIGIN
Query Match 55.3%; Score 466.2; DB 13; Length 905;
Best Local Similarity 69.2%; Pred. No. 4.2e-76;
Matches 584; Conservative 0; Mismatches 252; Indels 8; Gaps 3;
Qy 8 AAAANTTATTTNAATTTCTATTTAANCNTCTCCNCAANCATTATTTNACCCCTATNNGN 67
Db 9 AAAAGAAATCTAATTTTATTAACATTAATCTTCCCAAGCATTTATTTATCCTATATCT 68
Qy 68 CNGCAGNTTANAAAANTACCTTNTNTTAAAAAACTNGGAAAAAATAATNGCAA 127
Db 69 CACTGAATTTTAAAGAAATACATTAATGATTAAGAAAACTAGAAAAAAGATTAATGCAG 128
Qy 128 TANTTAACCTTCTGAAAANGAAATTTNTACCAANGACNCAANCTNTTAATTNGA 187
Db 129 TAAATTAACCTTACATGAAAAAGAAAAATTATTAACAAGAGACTGAGAACGTTATAAATTGA 188
Qy 188 ANTNAATTTATANTNGAAANGCNCNGAAACCAANCTTNAATGTCCTCAATTATCTTNA 247
Db 189 AATGAGATTATTAATTGAAAACTGCATCTGAAAGCAACTTATTTGTTCAATTATCTTAA 248
Qy 248 ANGAGGNNNTTANNACTAATNCCNGATTTTCCAAATANGAANCCNNNTTAAANNTN 307
Db 249 ATGATGCTGTTTATGACTTAATACACTGATTTTTCATTAAGAAAAACCATGTTAAATA 308
Qy 308 TTTNATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAATTTGATG 367
Db 309 TTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGAACATATTTCTTATTTGATTT 368
Qy 368 GGGAAAAAATNCNGTTCNNATACCCNNGAANNGCAANNTTTTAAATTTTAAACCCCC 427
Db 369 GGGAGAAAAATACGTGTTCTGATGACATGAATGCAAAAATTTTGAATTTTAAATCTCAC 428
Qy 428 TANTTTTAAANCTATNGAAANTNGATTANNAGCTTGAATTGC-CAACCTANTTNCNG 486
Db 429 TAAATTTAAGAACTATGAGAAATGATTAAATGACATGAAGTGACACACACTAATTAATG 488
Qy 487 GCCACNGTGGCNGTNGTNTTCTTACTTANTCCCCCAAGGAAANNCCTTAANCNGAAN 546
Db 489 GCCAGCTGTTGCATGTTGTTCTTACTTAGTTCTCCCAAGGAAAACTCTTAAACTGAAT 548
Qy 547 CTCNCCAAATAAACCTTAANTATCCTTGTAACCAANCAAAACCTTTTNGTTTACN 606
Db 549 CTTGACGAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTTGTGTACA 608
Qy 607 TANTCCTTGGATTAAAGGCTCCCCCAATTNATCCNGAACCCANTTTTCCCCCNAACCA 666
Db 609 TAGTTCTTTGGATTTACTGTTCTTAATTTTATTTCTGAAACTTAATTTTACCACAGACCA 668
Qy 667 TANTTACCATTTTACCTTGTAAGGCNCAAGTNGTTGCAANTNCCGCAANCAAGTANTNT 726
Db 669 TAAATTAACATATTAATCTTGTATGACAGANTGTATGCAATTCGGAAGCAGTAGTATA 728
Qy 727 CCCNGG---CNCCTTCCCCGANCCTTGGAAGAAAAAGGATNGTGCCCCCTTAAAAA 783
Db 729 CCATCAGGCTCTATTATCCAGAGCATAGAAAAACAGATTGGTCAACCCCATTAAGAA 788
Qy 784 ACAACCTTCCCCCNC-----TTGGCCAGAGNNTTNTTCCCGCTTAATCCGAACAATAA 839
Db 789 ACAATCTATACACACCGCTATTGGCAATGTGTAGTTTAAAGTCTAAATCTGACATAAA 848
Qy 840 AAAG 843

Db 849 AATG 852

RESULT 9
LOCUS AI826287 770 bp mRNA linear EST 21-DEC-1999
DEFINITION wk33f07.x1 NCI CGAP Pr22 Homo sapiens CDNA clone IMAGE:2417221 3' similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (HUMAN); contains element MER28 repetitive element ;, mRNA sequence. AI826287

ACCESSION AI826287 GI:5446958
VERSION AI826287.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 770)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmerit-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1454 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 404.
Location/Qualifiers
1..770
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2417221"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr22"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 271 a 111 c 99 g 285 t 4 others

ORIGIN

Query Match 55.3%; Score 466; DB 9; Length 770;
Best Local Similarity 72.5%; Pred. No. 4.7e-76;
Matches 528; Conservative 0; Mismatches 199; Indels 1; Gaps 1;

QY 2 GCCAATAAANTTATTTTCTTAACTTAACNTCCCTCCNCAANCAATTATTNACCT 61
Db 25 GCCAGAAAAAGTATTATTCTTCTATTAAACATCTCTCAAGCATTAATTATCT 84

QY 62 ATNNCNCNGANTTTNANAANTACCTTNTNTTAAAAAAGCTNGAAAAAATAAT 121
Db 85 ATATCTCACTGAATTTTAAAGAAATAACATTAGATTAGAAAAAAGTAA 144

QY 122 NGCAATANTTAACCTTCTGAAAANGAAATTTTACCAANGAANGAANCNTNTA 181
Db 145 TGCAGATAATTAACCTTACATGAAAAAGAAATATTATAACAAGACTGAGAAGTTATA 204

QY 182 ATTNGAANTNAATTAATANTTNGAAANGCGCNCNGAACAACANTTNATGTCAATTA 241

Db 205 AATTGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTTATTGTTCAAATTA 264

QY 242 TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGGAANCCNNNTTA 301
Db 265 TTCTTAATGATGTTGTTTATGACTAATACACTGATTTTCAATAAGAAACCATGTTA 324

QY 302 AAANTTTTNNATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAATT 361
Db 325 AAAATATTTTATTTTAAAAATAAGCCTGTGTGCAAGCTCTGATCATATTTCTTAAATT 384

QY 362 GGATTGGGAAAAAATNCGTTCNNATACCCNNGAANNGCAANTTTTAAATTTTAA 421
Db 385 TGATTGGGAAGAAATACTGTCTTGATAGCATGAATGCAAAATTTTGAATTTTAA 444

QY 422 CCCCCCTANTTTTAAANCTAINGAAAAANTNGATTANNGACTGAATTGC-CAACCTAN 480
Db 445 TCTCACTAATTTTAAGAACTAATTGAGAAATTGATTAATGACATGAAGTGACACACTAA 504

QY 481 TTNCNGCCACCGNGTGGCNTNGTNTTCTTACTTANTTCCCCCAAGGAANNCTTAAN 540
Db 505 TTACTGGCCAGCTGTGGCAATTGTGTTCTTACTTAAGTTCTCCCAAGGAACCTCTTAA 564

QY 541 CNGAANTCCNCCAAAAATAACCTTAANTATCTTGTGTAACCAANCAAAACCTTTTNG 600
Db 565 CTGAATCTTACAGCAATAATCCTTAATATACCTTGTGAAGCAAAACAAAGCTNTNTTG 624

QY 601 TTACNTANTCCTTGGATTTAACGGGTCCCAATTNATCCNGAACCAANTTTTCCCCC 660
Db 625 TTACATAGTCTTGTGATTNTACTGTCTCTAATTTTATTTCTGAACCTCAATTTTACCC 684

QY 661 NAACCATANTTACCATTTTACCTTGTGTAAGCNCAGTNGTTGCANTNCCGCAANCACT 720
Db 685 AGACCATATTAACCATTTTAACCTNTGTAATGCACAGTGTATGCAATTTCCGCAAGCAGT 744

QY 721 ANTNTTCC 728
Db 745 AGTATACC 752

RESULT 10
LOCUS AI565988 793 bp mRNA linear EST 12-MAY-1999
DEFINITION tns2a05.x1 NCI CGAP Kid1 Homo sapiens CDNA clone IMAGE:2171984 3' similar to gb:552450 NONSPECIFIC LIPID-TRANSFER PROTEIN PRECURSOR (HUMAN);, mRNA sequence. AI565988

ACCESSION AI565988 GI:4524440
VERSION AI565988.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmerit-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert length: 654 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 458
POLYA=No.
Location/Qualifiers

source		1..793	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:2171984"	
		/lab_host="DH10B"	
		/clone_lib="NCI CGAP_Kid11"	
		/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; plasmid DNA from the normalized library NCI CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."	
BASE COUNT	291 a	122 c	110 g 266 t 4 others
ORIGIN			
Query Match 55.0%; Score 463.4; DB 9; Length 793;			
Best Local Similarity 71.0%; Pred. No. 1.4e-75;			
Matches 562; Conservative 0; Mismatches 224; Indels 5; Gaps 3;			
OY	1	GGCCAAAAAANTATTTNAAATTCCTATTANCNTCCCTCCNCAANCAATATTNACC	60
Db	3	GGCCAGAAAAATTATTTTATTTTCTATTAAACATTCCTCTCAAGCATTTATTATCC	62
OY	61	TATNCGNCNGANTTTNANAANTACCTTNTNTTAAAAAACCTNGAAAAAAATAA	120
Db	63	TATATCTCACTGAATTTTAAAGAAATAACATTAGATTAGAAAACTAGAAAAAGATAA	122
OY	121	TNGCAAATANTTAACCTTCTCTGAAANGAAATTTTNTACCAANGAAGCAANCTNTT	180
Db	123	ATGCAGATAATTAACCTTACATGAAGAAAAAATTATACAAAGAGCTGAGAACGTTAT	182
OY	181	AATTNGAANTNAAATTATANTNGAANGCGCNCNGAAACCAANCTNATGTCGAATT	240
Db	183	AAATTGAAATGAGATTATTAATTGAAAACTGCATCTGAAGCAACTTATTGTTCAATT	242
OY	241	ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAAATANGAANCCNNNTT	300
Db	243	ATTCTTAATGATGGTGTTTATGACTAATACACTGATTTTCAATAAGAAACCCATGTT	302
OY	301	AAAAANTTTTNTATTTAAAAATAACCCNGTNTCCAACCCCNNGATCANATTCCTTNA	360
Db	303	AAAAATATTTTATTATAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTAATT	362
OY	361	TGGATTGGGAAAAAATNCNGTTCNNATACCCNNGAANNCAAAATTTTAAATTTTA	420
Db	363	TTGATTTGGGAAAAAATACTGTTCTGATAGCATGAATGCAAAAATTTTAGATTTTA	422
OY	421	ACCCCTANTTTTAAANCTAATNGAAANTNGATTANNGACTGAATTGC-CAACCTA	479
Db	423	ATCTCACTAATTTTAAGACTAATGAGAAATGATTATGACATGAAGTGACACACTA	482
OY	480	NTNCGGCCACCGNGTGGCCTNGTNTTCTTACTTANTCCCCCAAGGAANNCTTAA	539
Db	483	ATTACTGGCCAGCTGTTGGCATTTGTCTTCTTACTTAGTTCTCCCAAGGAAAACTTAN	542
OY	540	NCNGAANTCCNCCAAATAACCTTAANTATCCTTGTTAACCAANCAAAACCTTTTN	599
Db	543	ACTGAATCTTCAGCAGAAATATCTTAATATACCTTGTAGCAAAACAAAGCCTTTTG	602
OY	600	GTTTACNTANTCCTTGGATTTAACGGGTCCCAATTNATTCNGAACCCANTTTTCCC	659
Db	603	G-TTACATAAGTCTTTGGATTNTACTGCTCCTAANTTTATCTGAAACTCAATTNTACC	661
OY	660	CNAACCATANTTACCATTTTACCTTGGTAAAGCNCAGTNGTTGCANTNCCGCAAAACAG	719
Db	662	CAGACCATATAATACCATATACTTGTATATGACAGATTGTATGCAATTTGCAAGCAG	721
OY	720	TANTNTTCCCNGG--CNCTTTCCTCCCGANCCTTGGAAAAAAGGATNGTCCCCCC	776

Db	722	TAGTATACCATCAGGCTCTATTCAACCAGAGCATAGAAAAAACGAGATTGGTCAACCC	781
OY	777	TTAAAAAACAA	787
Db	782	ATAAGAAACAA	792
RESULT 11			
BU627147			
LOCUS			
DEFINITION			
UI-H-FG0-bct-i-12-0-UI.s1 NCI CGAP_EN1_2 Homo sapiens cDNA clone			
UI-H-FG0-bct-i-12-0-UI 3', mRNA sequence.			
BU627147			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
Unpublished			
Contact: Robert Strausberg, Ph.D.			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: James Martin			
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa			
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa			
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa			
Clone Distribution: Clone distribution information can be obtained			
from Dr. M. Bento Soares, bento-soares@uiowa.edu			
The following repetitive elements were found in this cDNA			
sequence: 1-43, >AT_rich#low_complexity (matched complement)			
309-335, >AT_rich#low_complexity (matched complement)			
Seq primer: M13 FORWARD			
POLYA=Yes.			
FEATURES			
Location/Qualifiers			
source			
1..762			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="UI-H-FG0-bct-i-12-0-UI"			
/cissue_type="Enchondroma cell line"			
/dev_stage="Adult"			
/lab_host="DH10B (Life Technologies)"			
/clone_lib="NCI CGAP_EN1_2"			
/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a			
modified polylinker; Site 1: EcoR I; Site 2: Not I;			
NCI CGAP_EN1_2 is a cDNA library containing the following			
cissue(s): Enchondroma cell line (2 cell lines). The			
library was constructed according to Bonaldo, Lennon and			
Soares, Genome Research, 6:791-806, 1996. First strand			
cDNA synthesis was primed with an oligo-dT primer			
containing a Not I site. Double stranded cDNA was ligated			
to an EcoR I adaptor, digested with Not I, and cloned			
directionally into pT7T3-Pac vector. The oligonucleotide			
used to prime the synthesis of first-strand cDNA contains			
a library tag sequence that is located between the Not I			
site and the (dT)18 tail. The sequence tag for this			
library is CCGGTCACTC. The cell lines was provided by Dr			
James Martin from University of Iowa.			
TAG_LIB=UI-H-FG0			
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)			
TAG_SEQ=CGGTCACTC"			
BASE COUNT	269 a	109 c	97 g 286 t 1 others
ORIGIN			
Query Match 54.8%; Score 462.2; DB 13; Length 762;			
Best Local Similarity 72.7%; Pred. No. 2.4e-75;			
Matches 537; Conservative 0; Mismatches 200; Indels 2; Gaps 2;			

OY		9	AAANTTATTTNAATTTCTATTAAACNTCCCTCCNCAANAATTAATTNACCCTAATNNCNC 	68
Db		19	AAAGTATTTTAAATTTCTATTAACAATTCTTCTCAAAGCATTAFTTTATCCTATATCTC 	78
OY		69	NCNGANTTTNANAAAATAACCTTTNNNTTAAAAAACCTNGGAAAAAAATAATNGCAAT 	128
Db		79	ACTGAATTTTAAAGAATAACATAGATATTAGAAAAAACTAGAAAAAAGATAAATGCAGAT 	138
OY		129	ANTTAACTTNTCTGAAAAANGAAAAATTNTTACCAANGACNCAANCNTTNTAATTNGAA 	188
Db		139	AATTTAACTTACATGAAAAAGAAAAATTAATACAAAGAGACTGAGAACGTTATAAATTGAA 	198
OY		189	NTNAAATTAATNTNGAAANCGCANNNGAAACCAANCCTTNA TGTC CAATTATCTTAA 	248
Db		199	ATGAGATTATAATTGAAAACTGCATCTGAAAGCAAACTTAATTGTCAATTATCTTAA 	258
OY		249	NGAGGNNTTTNANNACTAATNCCCNGAATTTTCCAAATANGAANCCCNNTTAAAAANTNT 	308
Db		259	TGATGGTGTTTTATGACTAATAACA CTGATTTTCCAATAGGAAAA C CATGTTAAAAATAT 	318
OY		309	TTTNAATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAATTTGGATTGG 	368
Db		319	TTTTATTTTAAAAATAAGCCTGTGTTC AAGCTCTGATCATATTTCTTTATTTGATTTG 	378
OY		369	GGAATAAAATNCGTTCNNNTAACNNGAANNGCAAANTTTTAAATTTTAAACCCCT 	428
Db		379	GGAAGAAATACTGTTCTGATAGCATGAATGCAAAATTTTAGATTTTAAATCTCACT 	438
OY		429	ANTTTTAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC -CAACCTANTTNCNG 	487
Db		439	AATTTTAAAGACTATTGAGAAATTTGATTAATGACATGAAGTGCAACA C ACTAATTA CTGG 	498
OY		488	CCACCGNTGGC NTNGTTCCTTA CT TANTCCCCCAAGAAAA NNCTTAA NCNGAANC 	547
Db		499	CCAGCTGTGGCATGTGTTCCTTA CT T AGTTCTCCCAAGAAAAA CTTTAAACTGAA TC 	558
OY		548	TCCNCCAAATAAC CCTTA ATATCTTG GTGAACCAAA NCAAAACCTTTTNGTTA CNT 	607
Db		559	TTCA GCA GAAT AAT CCTTA ATA TACTTTG TAAGCAAA ACAAAGCTTTTGTGTTACAT 	618
OY		608	ANTCCTTGGATTAA CGGGTCCCAATTT NATCCNGAACCCANTTTTCCCCNAA CCAT 	667
Db		619	AGTCTTTGGATT TTA CTGTTCCTAATTTTATTTCTGAAC TCAATTTT ACCCAGACCAA 	678
OY		668	ANTTACCATT TTACCTTG GTAAGCNCAGTNGTTGC ANTNCCGCA AANCAGTANT TTC 	727
Db		679	TATTACCATATTA CT TTGTATATG CACAGTTGTATGCA -TTCCGCANAGCAGTAGTATAC 	737
OY		728	CCCNNGGCNCTTCCCCCGA 746 	
Db		738	ATCAGGCTCTATTCA CCCA 756 	
RESULT 12				
BE738457 824 bp mRNA linear EST 15-SEP-2000				
LOCUS 601572489T1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:3839159 3',				
DEFINITION mRNA sequence.				
ACCESSION BE738457				
VERSION BE738457.1 GI:10152449				
KEYWORDS EST.				
SOURCE Homo sapiens (human)				
ORGANISM Homo sapiens				
REFERENCE				
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.				
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished				
Contact: Robert Strausberg, Ph.D.				
Email: cgapbs-r@mail.nih.gov				
Tissue Procurement: ATCC				
CDNA Library Preparation: CLONETECH Laboratories, Inc.				

Query Match	Best Local Similarity	Score	DB	Length	824;
Matches	565;	Conservative	0;	Mismatches	227;
			Indels	Gaps	3
51	TAATTNACCCCTATNNCNCNCNGANTTTNANAANTACCTTNNNTNTTAAAAA	54.6%;	460;	10;	824;
3	TAATTTATCCTAGATCTTACTGAATTTTGAATAATACATTAGATTAGAAAA	70.8%;	Pred. No. 5.9e-75;		
111	AAAAAATAATNGCAATANTTAACCTTCTGAAANGAAATTTTACCAANGACNG				
63	AAAAAGATAAATGCAGATAATTAACCTTACATGAAAAAGAAATTTAACAAGACTG				
171	AAANCNTNTAATNGAANTNAAATTAANTNGAANGCGCNCNGAACAACACTTNA				
123	AGACGTTATAAATGGAATGAGATTATAATTGAAACTGCATCTGAAGCAACTT				
231	TGGTCCAATTATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGA				
183	TTGTTCAATTATCTTAATGATGGTGTTTTATGACTAATACATGATTTTCAATAAGGA				
291	ANCCNNNTTAAAAANTTTTNTAATTTAAAAATAACCCNGTNTCCAACCCNGATCANAT				
243	AACCCATGTTAAAAATATTTTNTAATTTAAAAAATAAGCCCTGTGTTCAAGCTTGATCATAT				
351	TCCTTNTAATTTGGATTGGGAAAAAATNCGTTCNNATACCCNNGAANGCAAAANTTT				
303	TTCTTTAATTTGATTGGGAGAAATACTGTTTCTGATAGCATGAATGCAAAATTTT				
411	TAAATTTTAAACCCCTANTTTTAAAAANCTATNGAAAAANTNGATTANNGACTTGAATTG				
363	TAGATTTTAAATCTCACTAATTTTAAAGAACTATTGAGAAATGATTAATGACATGAAGTG				
471	C-CAACCTANTTNCNGGCCACCGTGGGCGTNGTNTTCTTACTTANTCCCCCAAGGA				
423	CACAACACTAATTACTGGCCAGCTGTGGCATTTGTGTTCTTACTTACTTCTCCCAAGGA				
530	AAANCCCTTAANCNGAANTCCNCCAAAAATAACCCCTTAANTATCCTTGGTAACCAANCAA				
483	AAACTCTTAACCTGAATCTTGACGACGAATAATCCTTAATAATATACTTTGTAAGCAAAACAA				
590	AACTTTTNGTTTACNTANTCTTGGGATTTAACGGGTCCCAATTNATCCNGAACCC				
543	AAGCTTTTGTGTACATAGTTCTTGGATTTTACTGTTCTCTAATTTATTTGAAACTC				

REFERENCE 1 (bases 1 to 822)
AUTHORS Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA HTC clones
JOURNAL Unpublished
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..822
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HTCCAD08"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="HTC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 292 a 133 c 115 g 277 t 5 others
ORIGIN

Query Match
Best Local Similarity 54.4%; Score 458.2; DB 9; Length 822;
Matches 552; Conservatve 0; Mismatches 216; Indels 5; Gaps 3;

QY 16 TTTNAAATTCCTATTAAACNTCCCTCCNCAAAACATTATTNACCTATNMCNCGANT 75
Db 16 TTTTAAATTTCTATTAAACATCTCTCAAGCATTATTTCCTATATCTCAGTGAAT 75
QY 76 TTNANAAANTACCTTTNNTNTTAAAAACCTNGGAAAAAATAATNGCAATANTTAAC 135
Db 76 TTTAAGAAATAACATTAGTATTAAGAAAACTAGGAAAAAAGATAAATGCAGATAATTA 135
QY 136 CTTCCTGAAANGAAATTTNTACCAANGACGACGAAANCNTNTAATTNGAANTNAAT 195
Db 136 CTTCACATGAAAAAGGAAAAATTATTAACAAGGACTGAGAACGTTATAAATTGAATGACAT 195
QY 196 TATANTTNGAAGGCGGCGNCCGAAACCAANTCTNATGGTCCAATATCTTNAANGAGG 255
Db 196 TATTAATTTGAAAACTGCATCTGAAGCAAACTTATGTTCATTAATTTCTTAATGATGT 255
QY 256 NTTTTNANNACTAATNCCCGATTTTCCAATANGAANGCCNNTTAAANTNTTTNAT 315
Db 256 GTTTTATGACTAATACACTGATTTTCAATAAGGAACCAATGTTAAAAATATTTTAT 315
QY 316 TTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNATTTGATGGGAAAA 375
Db 316 TTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTATTTTATTTGATTTGGAGAA 375
QY 376 AATNCGTTCNNATACCCNNGAANNCAANTTTTAAATTTTAAACCCCTANTTTTA 435
Db 376 AATACTGTTTCTGATAGCATGAATGCAAAATTTTAGATTTTAACTCTCACTAATTTTA 435
QY 436 AAANCTATNGAAAAANTNGATNNGACTGAATTGC-CAACCTANTTNCNGGCCACNG 494
Db 436 AGAACTATTGAGAAATGATTAATGACATGAAGTGACACACACTAATTACTGGCCAGCTG 495
QY 495 TGGGCGTNGTNTTCCTTACTTANTTCCCCCAAGGAAANNCCTTAANCNGAANTCCNCA 554
Db 496 TTGGCATTTGTCTTCTTACTTATGTTCTCCCAAGGAAAACTCTTAAACTGAATCTTGCA 555
QY 555 AAATAACCTTAANTATCTTGTGTAACCAANCAAAACCTTTTNGTTACNTANTCCTT 614

Db 556 GAATTAATCCTTAATATACTNTGTAAAGCAAAACAAAGCTTTTGTGTTACATAGTTCTT 615
QY 615 GGGATTTAAACGGGTCCCAATTTNATCCNGAACCCANTTTTCCCCNAAACATANTTACC 674
Db 616 TGGATTTNACTGCT-CTTAATTTTATTTCTGAAACTCAATTTACCCAGACCATTAATACC 674
QY 675 ATTTTACCTTGTGAAGGCGACGTNGTTTGACANTNCCGCAANACAGTANTNTTCCCGNG- 733
Db 675 ATATTAACNTNTGTAATGACAGANTGTATGCACCTCCGACANAGCAGTAGATATACCATCAGG 734
QY 734 --CNCCTTCCCGGANCCTTTGGGAAAAACGGATNGTCCCGCTTAAAAAA 784
Db 735 CTCTATTACCCAGACGATAGGAAAAAACAGAGATTGTACACACCCCATTAAGAA 787

RESULT 15
BQ045161
LOCUS 738 bp mRNA linear EST 21-FEB-2003
DEFINITION UI-CF-EN1-aei-a-17-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-aei-a-17-0-UI 3', mRNA sequence.
ACCESSION BQ045161
VERSION BQ045161.1 GI:19796248
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 738)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).
The following repetitive elements were found in this cDNA sequence: 316-342, >AT rich#Low_complexity (matched compliment)
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES
source
1..738
Location/Qualifiers
1..738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-aei-a-17-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not

I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT.

TAG_LIB=UI-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTGCTCAGGT"

BASE COUNT 260 a 102 c 97 g 278 t 1 others
ORIGIN

Query Match 54.0%; Score 455.6; DB 12; Length 738;
Best Local Similarity 72.9%; Pred.No. 3.9e-74;
Matches 517; Conservative 0; Mismatches 191; Indels 1; Gaps 1;

```
OY 2 GCCAAAAAANTATTATTTTAAATTTCTATTAAACNTCTCCNCAANCAATTATTNACCCCT 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 GCCAGAAAAAGTTATTTTAAATTTCTATTAAACATCTCTCTCAAGCATTTATTATCCT 78

OY 62 ATNNCNCNGANTTTNANAANTACCTTTNNNTTAAAAACCTNGGAAAAAATAAT 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 ATATCTCACTGAATTTTAAAGAAATACATTAGATTAGAAAACTAGAAAAAGATAAAA 138

OY 122 NGCAATANTTAACCTTCTGAAAAANGAAATTTTACCAANGAGCNGAANCNTNTA 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 TGCAGATTAATTAACCTTACATGAAAAAGAAAAATTAACAAGAGACTGAGAACGTTATA 198

OY 182 ATTNGAANTNAATTAATTANTTNGAAANGCGCNCNCAACCAACTTNATGGTCCAATTA 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 AATTGAATGAGATTATTAATTGAAAACTGCATCTGAAAGCAAACTTATTGTTCAATTA 258

OY 242 TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTTA 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 TTCTTAATGATGTTGTTTATGACTAATACACTGATTTTCAATAAGAAACCATGTTA 318

OY 302 AAANTNTTTNATTTTAAAAATAACCCNGTNTCCAAACCNGATCANATTCCTTNATTT 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 AAATATTTTATTTTAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTTATTT 378

OY 362 GGATTGGGAAAAAATNCNGTTCNNATACCNNGAANNCAANTTTTAAATTTTAA 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 TGATTTGGGAGAAAAATACTGTTCTGATAGCATGAATAAGCAAAATTTTAGATTTTAA 438

OY 422 CCCCCCTANTTTTAAAAAANCTAATNGAAAANTNGATTANNGACTGAATTGC-CAACCTAN 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 TCTCACTAATTTTAAGAACTATTGAGAAATGATTAAATGACATGAAGTGACACAACCTAA 498

OY 481 TTNCNGGCCACCCNGTGGCNTNGTNTTCCCTTACTTANTCCCCCAAGAAANNCTTAAN 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 TTAAGCCAGCTGTGGCATTTGTTTCTTACTTAGTTCTCCCAAGAAACCTCTAAA 558

OY 541 CNGAANTCCNCAAAATAACCTTAANTATCCTTGTTAAACCAANCAAAACCTTTTNG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 CTGAATCTTCAGCAGATAATCCTTAATATACCTTGAGCAAAACAAAGCTTTTGTG 618

OY 601 TTTACNTANTCCTTGGATTAAAGGGTCCCCCAATTNATCCNGAACCCANTTTCCCCC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 TTTACATAGTTCTTGGATTTTACTGTCTCTAATTTTATTCTGAAACTCAATTTTACCCC 678

OY 661 NAACCATANTTACCATTTAACCTTGTTAAGGCNCAGTNGTTGCANTNC 709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 AGACCATGATTACCATTAATTAACCTTGTGATGACAGTGTATGATTCATTCC 727
```

Search completed: November 27, 2003, 12:28:20
Job time : 2274.04 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 05:01:24 ; Search time 277.008 Seconds
(without alignments)
8215.019 Million cell updates/sec

Title: US-09-835-992A-22

Perfect score: 843

Sequence: 1 ggcacaaaanaattactttna.....taatccgaacatacaaaag 843

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N_Geneseq_19Jun03.*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	725	86.0	843	20	AAX40097
2	503.4	59.7	2663	22	AAH57501
3	436.2	51.7	714	20	AAX40094
4	411.8	48.8	687	20	AAX40095
5	313.6	37.2	486	24	ABV87267
6	216.2	25.6	351	16	AAT19975
7	169.6	20.1	262	22	AAH57293
8	110	13.0	173	25	ABX83093

9	84.6	10.0	397	25	ABX40797	Bovine EST associa
10	78.2	9.3	449	25	ABX48095	Bovine EST associa
11	69.6	8.3	397	25	ABX47851	Bovine EST associa
12	69.4	8.2	263	15	AAQ76489	Human genome fragm
13	69.4	8.2	340	23	ABV60916	Human prostate exp
14	61.8	7.3	18133	24	ABK40018	Human chemically p
15	61.8	7.3	18133	24	ABL32941	Human immune syste
16	55.2	6.5	3053	23	ABV21847	Human prostate exp
17	55.2	6.5	3053	23	ABV27671	Human prostate exp
18	53.2	6.3	2152	24	ABI99762	Mouse ischaemic co
19	51.8	6.1	18218	24	ABL33949	Human immune syste
20	51.4	6.1	19576	24	ABL70576	Chemically treated
21	51.4	6.1	19576	24	AAS61259	Human immune syste
22	51.2	6.1	18434	24	ABL34006	Human immune syste
23	50.6	6.0	5452	24	ABL33149	Human immune syste
24	50.6	6.0	7351	24	ABL32029	Human immune syste
25	50	5.9	883	22	AAL15210	Human breast cance
26	50	5.9	8663	22	AAS46711	Tumour suppressor
27	49.6	5.9	9515	22	AAS45452	Chemically pretrea
28	49.6	5.9	9515	24	ABN80244	Human chemically m
29	49.6	5.9	9515	24	ABX28307	DNA transcription
30	49.4	5.9	9707	24	ABL33420	Human immune syste
31	49	5.8	19131	22	AAS46717	Tumour suppressor
32	48.8	5.8	11209	24	ABN80136	Human chemically m
33	48.4	5.7	7201	24	ABL32336	Human immune syste
34	48.4	5.7	15667	24	ABL34146	Human immune syste
35	48.4	5.7	17491	24	ABL34575	Human metastasis a
36	48.2	5.7	5338	24	ABL70330	Chemically treated
37	48.2	5.7	5338	24	AAS61281	Human gene regulat
38	48.2	5.7	5338	24	ABK31379	Signal transductio
39	48.2	5.7	6923	24	ABL70275	Chemically treated
40	48.2	5.7	6923	24	AAS61198	Human gene regulat
41	48.2	5.7	6923	24	ABK31296	Signal transductio
42	48.2	5.7	7049	24	ABL54306	Chemically treated
43	48.2	5.7	7049	24	ABL32157	Human immune syste
44	48	5.7	7781	24	ABL33169	Human immune syste
45	47.8	5.7	37515	24	ABQ66998	Human angiogenesis

ALIGNMENTS

RESULT 1	
ID	AAX40097 standard; DNA; 843 BP.
AC	AAX40097;
XX	02-JUL-1999 (first entry)
DT	
XX	Gastric cancer associated gene.
DE	
XX	Cancer associated antigen; diagnosis; research; treatment; human;
KW	breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW	prostate cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO9904265-A2.
XX	
PD	28-JAN-1999.
XX	
PF	15-JUL-1998; 98WO-US14679.
XX	
PR	22-JUN-1998; 98US-0102322.
PR	17-JUL-1997; 97US-0896164.
PR	10-OCT-1997; 97US-0061599.
PR	10-OCT-1997; 97US-0061765.
PR	10-OCT-1997; 97US-0948705.
PR	11-OCT-1997; 97GB-0021697.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.
XX	

PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert B;
PI Tureci O;
XX
DR WPI; 1999-132448/11.
XX
XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
XX
XX
PS Claim 67; Page 696; 787pp; English.
XX
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
SQ Sequence 843 BP; 237 A; 184 C; 89 G; 215 T; 118 other;

Query Match 86.0%; Score 725; DB 20; Length 843;
Best Local Similarity 100.0%; Pred. No. 2.1e-162;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAAAAAAANTTATTNAAATTCCTATTAAACCTCCGCAAAACATTTATTTNACCC 60
DB 1 GGGCAAAAAAANTTATTNAAATTCCTATTAAACCTCCGCAAAACATTTATTTNACCC 60
QY 61 TATNNCNCNGANTTTNANAANAATCCTTTNTNTTAAAAAACCCTNGGAAAAAAATTA 120
DB 61 TATNNCNCNGANTTTNANAANAATCCTTTNTNTTAAAAAACCCTNGGAAAAAAATTA 120
QY 121 TNGCAATANTTAACCTTCTTGAAANGAAATTTNTACCAANGAAGCAANCTNTT 180
DB 121 TNGCAATANTTAACCTTCTTGAAANGAAATTTNTACCAANGAAGCAANCTNTT 180
QY 181 AATTNGAANTNAAATTATANTTNGAANGCGGNCNGAAGCAANCTNATGTCCAAAT 240
DB 181 AATTNGAANTNAAATTATANTTNGAANGCGGNCNGAAGCAANCTNATGTCCAAAT 240
QY 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGGAANCCNNNTT 300
DB 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGGAANCCNNNTT 300
QY 301 AAAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGCATCANATTCCTTNAAT 360
DB 301 AAAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGCATCANATTCCTTNAAT 360
QY 361 TGGATTGGGAAAAAATNCGTTCCNNATACCCNNGAANGCAANTTTTAAATTTTA 420
DB 361 TGGATTGGGAAAAAATNCGTTCCNNATACCCNNGAANGCAANTTTTAAATTTTA 420
QY 421 ACCCCCTANTTTTAAAAANCTAANGAAANTNGATTANNGACTGAATTGCCAACCCCTAN 480
DB 421 ACCCCCTANTTTTAAAAANCTAANGAAANTNGATTANNGACTGAATTGCCAACCCCTAN 480
QY 481 TTNCGGCCACCCNGTGGCGNTNGTTCCTTACTTANTCCCCCAAGAAANNCCTTAN 540
DB 481 TTNCGGCCACCCNGTGGCGNTNGTTCCTTACTTANTCCCCCAAGAAANNCCTTAN 540
QY 541 CNGAANCTCCNCCAAATAACCCCTTAANTATCCTTGTTACCAAAACCTTTTNG 600
DB 541 CNGAANCTCCNCCAAATAACCCCTTAANTATCCTTGTTACCAAAACCTTTTNG 600

QY 601 TTACNTANTCCTTGGATTAAACGGGTCCCAATTNATCCNGAACCCANTTTTCCCCC 660
DB 601 TTACNTANTCCTTGGATTAAACGGGTCCCAATTNATCCNGAACCCANTTTTCCCCC 660
QY 661 NAACCATANTTACCATTTTACCTTGTAAGGCNCAGTNGTTGCANTNCCGAANCACT 720
DB 661 NAACCATANTTACCATTTTACCTTGTAAGGCNCAGTNGTTGCANTNCCGAANCACT 720
QY 721 ANTNTTCCCGGCGNCTTTCCCGGAGAACCGGATNGTCCCGCTTAA 780
DB 721 ANTNTTCCCGGCGNCTTTCCCGGAGAACCGGATNGTCCCGCTTAA 780
QY 781 AAACAACCTTCCCGGAGAACCGGATNGTCCCGCTTAA 840
DB 781 AAACAACCTTCCCGGAGAACCGGATNGTCCCGCTTAA 840
QY 841 AAG 843
DB 841 AAG 843

RESULT 2
AAH57501/c
ID AAH57501 standard; cDNA; 2663 BP.
XX
AC AAH57501;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human liver cell specific cDNA sequence SEQ ID NO:341.
XX
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS Homo sapiens.
XX
PN WO200132927-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000WO-US30396.
XX
PR 04-NOV-1999; 99US-0163508.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sornasse T, Seilhamer JJ, Watson GA;
XX
DR WPI; 2001-291057/30.
XX
PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
XX
XX
PS Claim 1; Page 258-259; 327pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytosstatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a

CC cancer, immunopathology or neuropathology.

XX Sequence 2663 BP; 807 A; 472 C; 599 G; 785 T; 0 other;

Query Match 59.7%; Score 503.4; DB 22; Length 2663;

Best Local Similarity 71.7%; Pred. No. 9.5e-110;

Matches 609; Conservative 0; Mismatches 234; Indels 6; Gaps 3;

OY	1	GGCCAAAAAANTATTATTNAATTTCCTATTAA	NCNCTCCNCNCAAA	NCATTATTNA	CCC	60								
Db	2653	GGCCAGAAAAAGTTATTTAAATTTCTATTAA	CACTTCTCTCA	AAAGCATTATTTC	CAAGCATTATT	2594								
OY	61	TATNNCNCGANTTTNANAANTACCTTNTNTT	AAAAAACC	TNGAAAAAAATAA		120								
Db	2593	TATATCTACGTAATTTTAAGAAATAACATTAG	TATTAAGAAAACTAG	AAAAAAGATTA		2534								
OY	121	TNGCAATANTTAAACCTTNCCTTGAAANGAA	ATTNTTACCA	ANGAGACNGAA	NCNTNT	180								
Db	2533	ATGCAGATAATTAACTTACATGA	AAAAAGAA	AAATTATTA	CAAGAGCTGAGAACGTTAT	2474								
OY	181	AATNGAANTNAAATTATANTNGA	AAANCGC	NNCNGAA	CAANCTTNATGTC	CAATT	240							
Db	2473	AAATGGAATGAGATTATTAATTG	AAAACTGCATCTG	AAAGCA	ACTTATGTCTCA	ATT	2414							
OY	241	ATCCTNAANGAGGNNTTNANNACTA	ATNCCNGATTTTCCA	TANGA	ANCCNNNTT	300								
Db	2413	ATTCTTAATGATGTGTTTATGAC	TAAATACACTGATTTTCA	GAGAA	GGAAACCCATGTT	2354								
OY	301	AAANNTNTTTNATTTTAAAAATA	CCCNCGTNTCC	AAACCCNGAT	CANATTCCTTNA	TT	360							
Db	2353	AAAAATATTTTATTTTAAAAATA	AGCCCTGTGTTCA	AGCTCTGAT	CATATTTCTTTAT	TT	2294							
OY	361	TGGATTGGGAAAAAATNCNGTTC	NNATACCNNGA	ANNGCA	ANNTTTTAA	ATTTTA	420							
Db	2293	TTGATTTGGGAAAGAAATACTG	TTCTGATAG	CATGA	ATGCA	AAATTTT	TAGATTTTA	2234						
OY	421	ACCCCTANTTTTAAANCTA	TNGA	AAANTNG	ATTAN	NGACTTGA	ATTGC-CA	ACCTTA	479					
Db	2233	ATCTCACTAATTTTAAGAC	TATGAGAA	ATTGATTA	TATGACATGA	AGTGCACA	CACTTA	2174						
OY	480	NTTNCNGGCCACCGTGGC	NTNGTNTTCTTACT	TANTCCCC	CAAGGA	AAANCTTTAA	539							
Db	2173	ATTACTGGCCAGCTGTG	GCAATGTGTTCTTACT	TAGTTCTCC	CAAGGA	AAACCTCTTAA	2114							
OY	540	NCNGAANCTCCNCCAA	ATAACCTTA	ANTATCCTTGT	GAACCA	ANCAAAACCTTTTN	599							
Db	2113	ACTGAATCTTCAGCAG	AATATCTTAA	GCAAAATAC	CAAAAGC	TTT	2054							
OY	600	GTTTACNTANTCCTTGG	ATTAA	CGGGTCC	CAATTNAT	TCNGA	ACCANTTTTCCC	659						
Db	2053	GTTTACATAGTCTTGG	ATTAA	CTGTCT	CAATTTAT	TCGAA	ACTCAATTTTACC	1994						
OY	660	CNAACATANTTACCA	TTTACTTGT	GAAGC	NCAGTNGT	TGCANT	CCGCA	AAAG	719					
Db	1993	CAGACCATATTA	CCATATTA	CTTGT	ATATG	CACAGTGT	ATGCA	ATTCCG	CAAG	1934				
OY	720	TANTNTTCC--CC	NGCNC	TTCCCC	CGANC	CTTGG	GA	AAAA	CGGAT	NGTCCCC	CT	777		
Db	1933	TAGTATACCAT	CAGGCTCT	ATTTAC	CCCA	GAGCAT	AGAAAA	CA	GAGATT	GGT	CAC	ACCA	1874	
OY	778	T---AAAAA	CAACCTTCCCC	CNCCTT	TGGCC	CA	G	GNNTNT	TCCC	GT	TA	ATCC	GAC	834
Db	1873	TTAAGAA	CAAA	ATCTAT	CCAC	GCTAT	TTTGC	CAAT	GT	TA	GTTC	AG	1814	
OY	835	AAT	AAAAAG	843										
Db	1813	AAT	AAAAATG	1805										

RESULT 3
AAK40094
ID AAK40094 standard; DNA; 714 BP.

XX

AC AAK40094;
XX
DT 02-JUL-1999 (first entry)
XX
DE Gastric cancer associated gene.

XX
KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.

OS Homo sapiens.

XX
PN WO904265-A2.

XX
PD 28-JAN-1999.

XX
PF 15-JUL-1998; 98WO-US14679.

XX
PR 22-JUN-1998; 98US-0102322.

PR 17-JUL-1997; 97US-0896164.

PR 10-OCT-1997; 97US-0061599.

PR 10-OCT-1997; 97US-0061765.

PR 10-OCT-1997; 97US-0948705.

PR 11-OCT-1997; 97GB-0021697.

XX
PA (LUDW-) LUDWIG INST CANCER RES.

XX
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfeunshchuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;

XX
DR WPI; 1999-132448/11.

XX
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers

XX
PS Claim 67; Page 695; 787pp; English.

XX
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions

CC
CC characterised by the expression of various cancer associated antigens.

CC
CC The invention provides nucleic acid sequences and encoded polypeptides

CC
CC which are cancer associated antigen precursors expressed in human breast

CC
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and

XX
SQ Sequence 714 BP; 258 A; 100 C; 92 G; 260 T; 4 other;

Query Match 51.7%; Score 436.2; DB 20; Length 714;
Best Local Similarity 73.1%; Pred. No. 5.8e-94;
Matches 516; Conservative 0; Mismatches 187; Indels 3; Gaps 3;

OY	2	GCCAAAAA	ANTATTATTTAA	TTTCTCTATT	AA	NCNTCC	TCCNCA	AA	NCATTATT	NA	CCCT	61
Db	2	GCCAGAAAA	AGTTATTTTAA	TTTCTATTA	AA	CAATTC	TCTCA	AA	GCATTTAT	T	TCT	61

OY	62	ATN	CNCNCGANTTTN	ANA	ANTAC	CTTNTNTT	AAAA	AA	CC	TNG	AAAA	AAAA	TAAT	121
Db	62	ATAT	CTCAGTAATTT	TAAGAA	ATAACAT	TAGTAT	TGA	AAAA	AA	CTAG	AAAA	AA	GATTA	121

OY	122	NG	CAATANTTA	ACCTTNC	TGAA	AA	ANGAA	AA	TTNTT	TAC	CA	ANGA	CG	AA	NC	TNT	TA	181
Db	122	TG	CA	GAT	AT	TA	AA	CTT	ACAT	GA	AAAA	AA	GA	AA	AT	TA	TA	181

OY	182	AT	TNGA	ANTNAA	TATANT	TNGA	AN	CG	CNN	CGAA	AA	CA	AN	CT	TNAT	GT	CA	ATTA	241
----	-----	----	------	--------	--------	------	----	----	-----	------	----	----	----	----	------	----	----	------	-----

Db 182 AATTGAAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTATGTTCAATTA 241
QY 242 TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTTA 301
Db 242 TNCCTTAATGATGGTGTGTTTATGACTAATACTGATTTTCAAGAAAGAAACCCATGTTA 301
QY 302 AAANTNTTNNATTTTAAAAATAACCCNGTNTCCAAACCCCGATCANATTCCTTNAATT 361
Db 302 AAAATAATTTTATTTTAAAAATAAGCTGTGTTCAAGCTGTGATCATATTTCTTTATTT 361
QY 362 GGATTGGGGAATAAATNCNGTCCNNATACCCNNGAANNCAANTTTTAAATTTTAA 421
Db 362 TGATTGGGAANAATACTGTCTGATAGCATGAATGCAAAATTTTAGATTTTAA 421
QY 422 CCCCCCTANTTTTAAAAANTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCCCTAN 480
Db 422 TCTCACTAATTTTAAANAACTAATTGAGAAATTGATTAATGACATGAAAGTGACAAACACTTAA 481
QY 481 TTNCNGGCCAACCCNGTGGGCGTNGTNTCTTACTTANTCCCCCAA-GGAANNCCCTTAA 539
Db 482 TTACTGGCCAGCTGTGGCATGTGTTCTTACTTAGTCTCCCAAGGGAAGAACTCTTAA 541
QY 540 NCGAANCTCCNCCAAATAACCCCTAANTATCCTTGTTAACAAANCAAAACCTTTTNN 599
Db 542 ATTGAATCTTCAGCAGAAATAATCCTTAAATATACTTTGTAAGCAAAACAAAGCTTTT 601
QY 600 GTTTCACNTA-NTCCTTGGATTAAAGGGTCCCCCAATTNNATCCNGAACCCANTTTTCC 658
Db 602 GTTTCACATAGTCTTGGATTCTTGTCTGCTAATTTTATCTGAAGAACTCAATTTTACC 661
QY 659 CCNAACCATANTTACCATTTTACCTTGTGTAAGGCNCAGTNGTTGC 704
Db 662 CCAGACCATAATTACCATATTTACTTGTGTCACAGTGTGTTGC 707

RESULT 4
AAX40095
ID AAX40095 standard; DNA; 687 BP.

XX AAX40095;
AC
DT 02-JUL-1999 (first entry)
XX
DE Gastric cancer associated gene.
XX
KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer; ss.
XX
OS Homo sapiens.
XX
PN WO904265-A2.
XX
PD 28-JAN-1999.
XX
PE 15-JUL-1998; 98WO-US14679.
XX
PR 22-JUN-1998; 98US-0102322.
PR 17-JUL-1997; 97US-0896164.
PR 10-OCT-1997; 97US-0061599.
PR 10-OCT-1997; 97US-0061765.
PR 10-OCT-1997; 97US-0948705.
PR 11-OCT-1997; 97GB-0021697.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert B;
PI Tyreci O;
XX
DR WPI; 1999-132448/11.
XX

PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
PS Claim 67; Page 695; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
XX
SQ Sequence 687 BP; 242 A; 98 C; 85 G; 258 T; 4 other;

Query Match 48.8%; Score 411.8; DB 20; Length 687;
Best Local Similarity 71.8%; Pred. No. 3.5e-88;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

QY 15 ATTNAATTTCTATTAAACNTCTCCNCAANCAATTATTNACCCTATNCGNCGAN 74
Db 1 ATTTAATTTCTATTAACATTTCTCTCAAGCAATTATTTATCTATATCTCACTGAA 60
QY 75 TTTNANAANTACCTTNNNTTAAAAAAACCTNGAAAAAATAATNGCAATANTTAA 134
Db 61 TTTTANAATAATACATTAGTATTGAAAAAACTAGAAAAAAGATNAATGCAGTAATTAA 120
QY 135 CCTTNTCTGAAAAANGAAATTTNTACCAANGACNGAANCNTNTAATNGAANTNAAA 194
Db 121 ACTTACATGAAAAAGGAAATTTATAACAAGAGACTGAGAACGTTATAAATTGAAATGAGA 180
QY 195 TTATANTTNGAAAAACGCGNCCNGAAGAACCAACTTNATGTCCAATTATCTTNAANGAGG 254
Db 181 TTATAATTTGAAAACTGCATCTGAAAGCAAACTTATGTTCAATTATCTTAAATGATGG 240
QY 255 NNTTNANNACTAATNCCNGATTTTCCAAATANGAANCCNNNTTAAANNTTTNAT 314
Db 241 TGTTTATGACTAATFACATGATTTTTCATFAGAGAAACCATGTTAAATAATTTTAT 300
QY 315 TTTAAATAFACCCGNTTCCAAACCCCGATCANATTCCTTNAATTGGATGGGAAAA 374
Db 301 TTTAAAAATAAGCCGTGTGTTCAAGCTGTGATCATATTTCTTTATTTGATGGGAAGA 360
QY 375 AAATNCGTCCNNATACCCNNGAANNCAANTTTTAAATTTTAAACCCCTANTTTT 434
Db 361 AATACTGTTCTGATAGCATGAATGCAAAATTTTATGATTTTAACTCNCCTAATTTT 420
QY 435 AAAANCTATNGAAAAANTNGATTANNGACTTGAAATTGC-CAACCCCTANTNCGGCCACCN 493
Db 421 AAGAACTATTGAGAAATGATTAATGACATGAGTGACACAACACTAATTACTGGCCAGCT 480
QY 494 GTGGCGTNGTNTTCTTACTTANTCCCCCAAGAAANNCCCTTAANCGAANCTCCNCC 553
Db 481 GTTGCATGTGTCTTACTTAGTCTCCCAAGGAAAACTCTTAACTGAACTTCCAGC 540
QY 554 AAATAACCCCTAANTATCCTTGTAACCAANCAAAACCTTTTNGTTTACNTANTCCT 613
Db 541 NGATAACCTTAATATACTTTGTTAGCCAAAC--AAAACCTTTTGTGTTACATAAGTTCT 598
QY 614 TGGGATTTACGGGTCCCAATTNATCCNGAACCCANTTTTCCCCCNACCATANTTAC 673
Db 599 TTGATTTTACTGTTCTTAATTTATTTGAAAACTCCATTTTCCCGACACATAATTAC 658
QY 674 CAT-TTTACCTTGTTAAGGCNCAGTNGTT 701

DB 659 CCTATTTAACCTTGTTATGCACAGTTGTT 687

RESULT 5
ABV87267 standard; CDNA; 486 BP.

ID ABV87267
XX AC ABV87267,
XX DT 13-DEC-2002 (first entry)
XX DE Human colon cancer related CDNA SEQ ID NO 578.
XX KW Human; colon; cancer; cytostatic; tumour; gene therapy; vaccine; gene;
XX KB gb.
XX OS Homo sapiens.
XX PN WO200258534-A2.
XX PD 01-AUG-2002.
XX PF 19-NOV-2001; 2001WO-US43704.
XX PR 20-NOV-2000; 2000US-252222P.
PR 06-FEB-2001; 2001US-267011P.
PR 28-MAR-2001; 2001US-279670P.
PR 10-JUL-2001; 2001US-304037P.

PA (CORI-) CORIXA CORP.
PI Stolk JA, Xu J, Chenault RA, Meagher MJ, Secrist H, King GE;
XX WPI; 2002-608400/65.

DR New isolated tumor colon polynucleotide and polypeptide, useful for the
PT diagnosis, prevention and/or treatment of cancer, in particular colon
PT cancer -
PS Claim 1; SEQ ID NO 578; 266pp + Sequence Listing; English.

XX The invention relates to a human colon tumour expressed polynucleotide
CC (I) encoding a polypeptide (II, ABP67991-ABP67996) comprising: (i) any of
CC 2600 fully defined nucleotide sequences (ABV8669-ABV89289); (ii)
CC complements of (i); (iii) at least 20 contiguous residues of (i); (iv)
CC sequences that hybridize to (i), under moderately stringent conditions;
CC (v) sequences having at least 75% or 90% identity to (i); or (vi)
CC degenerate variants of (i). The compositions and methods of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC cancer, particularly colon cancer. (I) can be used in gene therapy and
CC (i) and (ii) are useful in pharmaceutical compositions such as vaccines.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 486 BP; 192 A; 55 C; 64 G; 175 T; 0 other;

[illegible]

RESULT 6	
AAT19975/c	
ID	AAT19975 standard; cDNA to mRNA; 351 BP.
XX	
AC	AAT19975;

DT	17-JUL-1996 (first entry)
XX	
DE	Human gene signature HUMGS01107.

KM Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KM human; cloning; mapping; non-biased library; diagnosis; detection;
KM cell typing; abnormal cell function; 88.

05 Homo sapiens.

PN W09514772-A1.

PD . 01-JUN-1995 .

PF 11-NOV-1994; 94WO-JP01916.

PR 12-NOV-1993; 93JP-0355504.

PA (MATS/) MATSUBARA K.

РА (ОКУБ/) ОКУБО К.

PI Matsubara K, Okubo K;

DR WPI; 1995-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 525; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(r) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of

CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

XX Sequence 351 BP; 117 A; 44 C; 44 G; 141 T; 5 other;

Query Match 25.6%; Score 216.2; DB 16; Length 351;
Best Local Similarity 73.6%; Pred. No. 6.8e-42;
Matches 256; Conservative 0; Mismatches 90; Indels 2; Gaps 2;

QY 1 GGCCAAAAAANTTATTNAAATTCCTTAATTAANCNTCCCTCCNCAANCAATTATTNA-CC 59

Db 348 GGCCAGAAAAAGTTATTTTAAATTCCTTAATTAACAATCTCCCAAGCATTATTANCC 289

QY 60 CTATTNCGNCGANTTTNANAAANTACCTTTNTNTTAAAAAACCCTNGAAAAAAATA 119

Db 288 CTATATCTCACTGAATTTTAAAGAAATACATTAGTATTAGAAAAAAGAAAAAGATA 229

QY 120 ATNGCAAAATANTTAACCTTNCCTGAAAGAAATTTNTACCAANGAANGAANCNTN 179

Db 228 AATGCAGATTAATTAACCTTACATGAAAAAGAAATATATAACAAGACTGGAACCTTA 169

QY 180 TAATNGAANTNAAATATATANTNGAAANGCGCNCNGAAGCAACCAANCTNATGTCCAAT 239

Db 168 TAAATGGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTATTGTTCAAT 109

QY 240 TATCCTNANGAGGGNNTTTNANNACTAATNCCNGATTTTCCAAATANGAANGAANCNNNT 299

Db 108 TATCTTAATGATGCTGTTTATGACTAATACACTGATTTTCAATTAAGGAACCACATGT 49

QY 300 TAAAAANTTTT-TNATTTAAAAATTAACCCNGTNTCCAACCCCGCATC 346

Db 48 TAAAAATATTNTNTNTTTAAAAAATAAGCCTGTGTTCAAGCTCTGATC 1

RESULT 7
AAH57293/c
AAH57293 standard; cDNA; 262 BP.

AC AAH57293;

DT 10-SEP-2001 (first entry)

DE Human liver specific cDNA sequence SEQ ID NO:133.

XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

OS Homo sapiens.

XX WO200132927-A2.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US30396.

XX 04-NOV-1999; 99US-0163508.

XX (INCY-) INCYTE GENOMICS INC.

XX Sornasse T, Seilhamer JJ, Watson GA;

XX WPI; 2001-291057/30.

PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -

PS Claim 1; Page 117; 327pp; English.

XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytosstatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.

XX Sequence 262 BP; 103 A; 38 C; 34 G; 87 T; 0 other;

Query Match 20.1%; Score 169.6; DB 22; Length 262;
Best Local Similarity 74.7%; Pred. No. 7.1e-31;
Matches 195; Conservative 0; Mismatches 65; Indels 1; Gaps 1;

QY 257 TTNANNACTAATNCCNGATTTTCCATANGAANGAANCCNNNTTAAANTNTTNAATTT 316

Db 262 TTTTATGACTAATACTACTGATTTTCAATAAGGAACCAATGTTAAAAATATTTTATTT 203

QY 317 TAAAAATPAACCCNGTNTCCAAACCCCGATCANATTCCTTNAATTTGGATTGGGAAAAAA 376

Db 202 TAAAAATPAAGCCTGTGTTCACAGCTCTGATCATATTTCTTTATTTGATTGGGAAGAAA 143

QY 377 ATNCGTTCNNATACCCNNGAANGCAAAANTTTTAAATTTTAAACCCCTTANTTTTAA 436

Db 142 ATACTGTTTCTGATAGCATGAATGCAAAATTTTGAATTTTAACTCATAATTTTAA 83

QY 437 AANCTATNGAAANTNGATTANNGACTTGAATTC-CAACCCCTANTTNCNGCCACCGT 495

Db 82 GAACATATTGAGAAATGATTAATGACATGAAGTGCAACAACATAATTACTGGCAGCTGT 23

QY 496 GGGCNTNGTNTTCTTACTTA 516

Db 22 TGGCATTTGTGTTCTTACTTA 2

RESULT 8
ABX83093/c
ABX83093 standard; cDNA; 173 BP.

AC ABX83093;

DT 24-APR-2003 (first entry)

DE Corn ear-derived polynucleotide (cpd) #1553.

XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022;

KW SATMON023; structural gene; functional gene; regulatory gene;
KW corn ear-specific profile; gene transcription; gene expression;
KW hybrid plant; desirable trait expression; plant breeding program;
KW inheritance; desired characteristic; growth; development;
KW disease resistance; environmental adaptability; quality; yield;
KW multigene trait; plant; gene; ss.

XX Zea mays.

XX US6476212-B1.

XX 05-NOV-2002.

XX 14-MAY-1999; 99US-0313294.

XX 26-MAY-1998; 98US-086722P.

PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
DR WPI; 2003-208840/20.
XX
PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with
PT growth, development
XX
PS Examples; SEQ ID No 1553; 390pp; English.
XX
CC The present invention relates to the isolation of corn ear-derived
CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries
CC SATMON022 and SATMON023. Some of the cpds uniquely identify structural,
CC functional, and regulatory genes of corn ear. The polynucleotides
CC sequences are useful for detecting cpds in a sample, for producing
CC a corn ear-specific profile of gene transcription, for detecting
CC altered gene expression in inbred or hybrid plants, and for screening
CC several molecules for specific binding to the polynucleotide. The cdps
CC are useful to identify, isolate, or extend identical or related
CC corn-ear nucleic acid sequences from DNA libraries, and in nucleic
CC acid amplification or hybridisation techniques to follow the
CC expression of desirable traits through plant breeding programs.
CC Preferably, the cdps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with
CC growth and development, disease resistance, environmental adaptability,
CC quality, and yield of corn. The cdps are also useful as molecular
CC markers for studying inheritance and multigene traits in a plant
CC breeding program. The cpds are useful for producing purified corn-ear
CC polypeptides by recombinant techniques. They are also useful in
CC diagnostic assays to detect or confirm conditions or diseases
CC associated with abnormal levels of cdp expression. ABX81541-ABX89140
CC represent corn ear-derived polynucleotides (cpds) of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/patseq/entry.html.
XX
SQ Sequence 173 BP; 54 A; 20 C; 19 G; 80 T; 0 other;
Query Match 13.0%; Score 110; DB 25; Length 173;
Best Local Similarity 71.3%; Pred. No. 8.6e-17;
Matches 122; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 8 AAAANTTATTNATTTCCCTATTAAACNTCCCNCAANCAATTATTNACCCTATNCCN 67
DB 172 AAAAAGTATTATTATTTCTATTAACATTTCTCTCAAGCATTTATTTATCTATATCT 113
QY 68 CNCNGANTTNANAAANTACCTTNNNTNTAAAAAACCTNGGAAAAAAATAATNGCAA 127
DB 112 CACTGAATTTTAAAGAAATAACATTAGTATTGAAAAAACTAGGAAAAAAGATAATGCAGA 53
QY 128 TANTTAACCTTCTGAAAAANGAAATTNTACCAANGAGCAAGCAANCNTT 178
DB 52 TAAATTAACTTACATGAAAAAGAAATTATTAACAAGGACTGAGAACGTT 2
RESULT 9
ABX40797
ID ABX40797 standard; cDNA; 397 BP.
XX
AC ABX40797;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #5962.
XX
KW Bovine; BS; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.

XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
XX WPI; 2003-110599/10.
DR
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2; SEQ ID No 5962; 245pp; English.
PS
XX
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a
XX second nucleic acid molecule comprising any of 1512 nucleotide
XX sequences, appearing as ABX34836-ABX49947, or complements of them.
XX Also included are; (1) a transformed cell having a nucleic acid
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-
XX translated sequence that functions in the cell to cause termination of
XX transcription and addition of polyadenylated ribonucleotides to a 3' end
XX of the mRNA molecule; and (2) determining a level or pattern of a
XX molecule in a bovine cell or tissue comprising: (a) incubating a marker
XX nucleic acid (comprising any of the 1512 nucleic acid sequences or its
XX complement or fragment) with a complementary nucleic acid molecule
XX obtained from the bovine cell or tissue, where hybridisation between the
XX marker nucleic acid and the complementary nucleic acid permits the
XX detection of the molecule; and (b) detecting the level or pattern of the
XX complementary nucleic acid, where the detection of the complementary
XX nucleic acid is predictive of the level or pattern of the molecule.
XX The LMFD nucleic acid is used for determining a level or pattern
XX of a molecule in a bovine cell or tissue. It is useful for genome
XX mapping, gene identification and analysis, cattle breeding, preparation
XX of constructs for use in cattle gene expression, or for genetically
XX improving cattle. The present sequence is one of the 1512 bovine
XX LMFD EST (expressed sequence tag) nucleic acids.
XX Note: The present sequence was not shown in the specification but
XX was obtained in electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139.
SQ Sequence 397 BP; 144 A; 75 C; 45 G; 133 T; 0 other;
Query Match 10.0%; Score 84.6; DB 25; Length 397;
Best Local Similarity 54.1%; Pred. No. 1.1e-10;
Matches 211; Conservative 0; Mismatches 167; Indels 12; Gaps 4;
QY 438 ANCTATNGAAAAANTNGATTANNGACTTGAATTGCCCAACCTANTTNCNGGCCACCGTGG 497
DB 18 ATCACTGAGAAATATATTAAGTCATTAAATGACACACATTAATTACTGTGACGTAATA 77
QY 498 GCNTNGTNTTCTTACTTANTTCCCCCAAGAAANNCTTAANCGAANGCTCCNCAAAA 557
DB 78 ATGATATTCTGTCTTCTCATTTCTCCCAAGAAACT-----TGAATTTCGGCAGAA 129
QY 558 TAACCCCTTAANTATCCTTGTAACCAANCAAAACCTTTTNGTTTACNTANTCCTTGGG 617
DB 130 TAACTCTCAATATACATTATTAAGCAAAATGAGAGCTTCT--GTTTACATACT-TTTTGT 186
QY 618 ATTTAAGGGGTCCCAATTNATCCNGAACCCANTTTTCCCCCNAAACCATANTTACCATT 677

Db 187 ATTTGCTATTCTTAAGTCTTATCTAAACTCAATTTTACCACCAACCATGATTACCATATA 246
QY 678 TTACCTTGGTAAGCNCAGTNGTTGCANTNCCGAAANCAGTANTNTTCC-CCNGGCNC 736
Db 247 TTAAGTTGTAATGCACAGTGTGTCACACTCAGCAAGCAGTAGTAACAACATCAGGTTTC 306
QY 737 TTTCCCGCCGANCCTTGGGAAAAACGGGATNGTCCCCCCTTAAAAACAACCTTCCCCC 796
Db 307 TATTACCCACGACACTTAGAAAAAACAGAAATTGATCACACCCCATTTAAAAACAATCTCAT 366
QY 797 NCCTTGGCCCGACGANNNTNTTCCCGTCTAA 826
Db 367 GCTATTTCCTCAATGTAGAGTTTCAATCTAA 396
RESULT 10
ABX48095/c
ID ABX48095 standard; cDNA; 449 BP.
XX
AC ABX48095;
XX
DT 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #13260.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle -
XX
XX
PS Claim 2; SEQ ID No 13260; 245pp; English.
XX
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.

CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.
XX
SQ Sequence 449 BP; 148 A; 70 C; 89 G; 142 T; 0 other;
Query Match 9.3%; Score 78.2; DB 25; Length 449;
Best Local Similarity 56.5%; Pred. No. 3.8e-09;
Matches 139; Conservative 0; Mismatches 106; Indels 1; Gaps 1;
QY 597 TTNGTTTACNTANTCCTTGGGATTTAACGGGTCGCCAATTNATCCNGAACCCANTTTTC 656
Db 430 TTCTGTTACCATACTTTTGTATTTTGTATTTCTAATCTTATTCTAAACTCAATTTTA 371
QY 657 CCCNAACCATANTTACCAATTTTACCTTGTAAGCNCAGTNGTTGCANTNCCGCAAN 716
Db 370 CCCAGACCATGATTACCATATTAAGTGTGTAATGCACAGTTGTTGCAACTCAGCAAG 311
QY 717 CAGTANTNTTCC-CCNGGCNCTTTCCCGGANCCTTGGGAAAAACGGGATNGTCCCCCC 775
Db 310 CAGTAGTAACCATCAGGTTCTATTTCACCCAGCACCTTAGAAAAACAAGAAATGATCACAC 251
QY 776 CTTAAAAACAACCTTCCCCCNCTTTGGCCCGCAGANNNTNTTCCCGTCTAAATCGAACA 835
Db 250 CCATTTAAAAACAACCTCATGCTATTTCCTCCCAATGTATAGTTTCAATCTAAGTTCTGACA 191
QY 836 ATAAAA 841
Db 190 ATGAAA 185
RESULT 11
ABX47851/c
ID ABX47851 standard; cDNA; 397 BP.
XX
AC ABX47851;
XX
DT 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #13016.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999; 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and

PT analysis, cattle breeding, or for genetically improving cattle -
XX
PS Claim 2; SEQ ID No 13016; 245bp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived
CC from cattle, and the LMPD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are; (1) a transformed cell having a nucleic acid
CC comprising an LMPD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMPD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMPD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.

XX
SQ Sequence 397 BP; 124 A; 57 C; 80 G; 136 T; 0 other;

Query Match 8.3%; Score 69.6; DB 25; Length 397;
Best Local Similarity 56.2%; Pred. No. 4e-07;
Matches 163; Conservative 0; Mismatches 123; Indels 4; Gaps 3;

QY 553 CAAATAACCCCTTAAATATCCTTGGTAACCAANCAAAACCTTTTNGTTACNTANTCC 612
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
397 CAGATAAGTCTTCAATATATCAATTATTAGCAAAATGAGAGCTTCT--GTTACATACTCA 340
QY 613 TTGGGATTTAAGGGTCCCAATTNATCCNGAACCCANTTTTCCCCNAAACCATANTTA 672
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 TT-GTATTCTGCTATTATTAACTATCTTCTAAAACTCAGCTTTTACCCCAACCATGATTA 281
QY 673 CCATTTTACCTTGGTAAGCAGTNGTTGCANTNCCGCAANCAGTANTNTTCC-CCN 731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
280 CCATATTAACTGTGTAATGACAGATGTAGCAACTCAGCAATGCTGTAGTAAACCATCA 221
QY 732 GGCNCTTCCCCCGANCTTGGGAAAAAAGCGATNGGTCCGCCCTTAAAAAACAACCTT 791
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 GGTATATATTCACCCAGCAATGATTAATAAAGAGACTTGATCAGACCATTTAAAAACAAC 161
QY 792 CCCCCNCTTTGGCCAGGNNNTTNTCCCGTCTAATCCGAACATAAAA 841
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 CTCATGCTATTTCCTCAATCTATAGTTTCAATCTAATCTTGACAAATGAAA 111

RESULT 12
AAQ76489/c
ID AAQ76489 standard; DNA; 263 BP.
XX
AC AAQ76489;
XX
DT 25-MAR-2003 (updated)
DT 23-SEP-1994 (first entry)
XX
DE Human genome fragment.
XX
KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
KW detection; homology; human; adrenal tissue; ds.
XX

OS Homo sapiens.
XX
PN WO9401548-A2.
XX
PD 20-JAN-1994.
XX
PF 13-JUL-1993; 93WO-GB01467.
XX
PR 13-JUL-1992; 92GB-0014857.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;
PI Sibson DR, Starkey M;
XX
DR WPI; 1994-035056/04.
XX
PT New nucleic acid fragment encoding gene products - can be used
PT for genetic analysis and mapping
XX
PS Claim 1; Page 64; 616pp; English.

CC Human nucleic acid fragments, isolated from brain adrenal tissue,
CC the placenta or bone marrow comprise any of: (A) a sequence
CC selected from (AAQ76401-Q77613), (B) an allelic variation of a
CC sequence as described in (A), or (C) a sequence complementary
CC to (A) or (B).
CC (Updated on 25-MAR-2003 to correct PN field.)
XX

SQ Sequence 263 BP; 97 A; 38 C; 45 G; 80 T; 3 other;

Query Match 8.2%; Score 69.4; DB 15; Length 263;
Best Local Similarity 60.7%; Pred. No. 4e-07;
Matches 153; Conservative 0; Mismatches 92; Indels 7; Gaps 4;

QY 463 TTGAATTGCCAACCCCTAATTNNGCCACCGNGTGGGNTNGTNTTCCTTACTTANTCCCC 522
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
263 TGGAGTACACAGCACTGATTAAGTCTGGCCAGCTGTGGCATTTGTGTTTC-TACTTACTTTTC 205
QY 523 CCAAGGAAANNCCTTAANGAANCTCCNCAAAATTAACCTTAANTATCCTTGTAACC 582
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
204 ACGAGGAAAACTCTTAAGTGAATCTTAGCAGAAATGCC--TTAATATACTTTGTAA-C 149
QY 583 AAANCAAAACCTTTTNGTTTACNTANTCCTTGGGATTTAAGGGTCCCAATTNATCC 642
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
148 AAAACAAAGCTTTTGTGTTTACATCGTCTTGTGATTTTATTTTATTTATTTCT 89
QY 643 NGAACCCANTTTTCCCCNAAACCATANTTACCATTTTACCTTGTAAGGCNCAAGTNTT 702
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 GAAACTCGATTTTACNNCAA--ACCATATTACTATTATTTGTAATGACACAGTTGAT 31
QY 703 GCANTNCCGCAA 714
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
30 GCAATTGCAAAA 19

RESULT 13
ABV60916
ID ABV60916 standard; cDNA; 340 BP.
XX
AC ABV60916;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 60907.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 06:06:02 ; Search time 317.882 Seconds
(without alignments)
8737.270 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 843
Sequence: 1 ggccaaaaaantcattttna.....taaatccgaacaataaaag 843

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	725	86.0	843	9	US-09-835-992A-22
2	505	59.9	1052	10	US-09-880-107-3296
3	436.2	51.7	714	9	US-09-835-992A-19
4	411.8	48.8	687	9	US-09-835-992A-20
5	318.8	37.8	506	10	US-09-796-692-7228
6	318.8	37.8	506	14	US-10-040-862-7228
7	313.6	37.2	486	10	US-09-998-598-578
8	182.2	21.6	510	14	US-10-102-524-161
9	182.2	21.6	510	14	US-10-102-524-188
10	182.2	21.6	510	14	US-10-102-524-645
11	84.6	10.0	397	10	US-09-960-352-5962
12	78.2	9.3	449	10	US-09-960-352-13260
13	69.6	8.3	397	10	US-09-960-352-13016
14	61.8	7.3	18133	12	US-10-311-455-914
15	51.8	6.1	18218	12	US-10-311-455-1922
16	51.2	6.1	18434	12	US-10-311-455-1979

C 17	50.6	6.0	5452	12	US-10-311-455-1122	Sequence 1122, Ap
C 18	50.6	6.0	7351	12	US-10-311-455-2	Sequence 2, Appli
C 19	50	5.9	960	14	US-10-198-846-6381	Sequence 6381, Ap
C 20	49.6	5.9	9515	12	US-10-240-453-181	Sequence 181, App
C 21	49.6	5.9	9515	14	US-10-239-676-159	Sequence 159, App
C 22	49.4	5.9	9707	12	US-10-311-455-1393	Sequence 1393, Ap
C 23	49.2	5.8	3673778	12	US-10-312-841-2	Sequence 2, Appli
C 24	48.4	5.7	7201	12	US-10-311-455-309	Sequence 309, App
C 25	48.4	5.7	15667	12	US-10-311-455-2119	Sequence 2119, Ap
C 26	48.4	5.7	17491	12	US-10-240-485-128	Sequence 128, App
C 27	48.2	5.7	7049	12	US-10-311-455-130	Sequence 130, App
C 28	48.2	5.7	7049	12	US-10-240-452-6	Sequence 6, Appli
C 29	48	5.7	7781	12	US-10-311-455-1142	Sequence 1142, Ap
C 30	48	5.7	3673778	12	US-10-312-841-1	Sequence 1, Appli
C 31	47.4	5.6	5864	12	US-10-311-455-1413	Sequence 1413, Ap
C 32	47.4	5.6	5864	12	US-10-240-452-61	Sequence 61, Appli
C 33	47.2	5.6	6103	12	US-10-311-455-2143	Sequence 2143, Ap
C 34	47.2	5.6	18011	12	US-10-311-455-7	Sequence 7, Appli
C 35	47	5.6	6577	12	US-10-311-455-1329	Sequence 1329, Ap
C 36	47	5.6	6823	12	US-10-311-455-1118	Sequence 1118, Ap
C 37	47	5.6	10132	12	US-10-311-455-2053	Sequence 2053, Ap
C 38	47	5.6	10132	12	US-10-240-452-71	Sequence 71, Appli
C 39	46.8	5.6	6494	12	US-10-311-455-1366	Sequence 1366, Ap
C 40	46.8	5.6	10189	12	US-10-311-455-2133	Sequence 2133, Ap
C 41	46.6	5.5	10855	12	US-10-311-455-1461	Sequence 1461, Ap
C 42	46.2	5.5	884	10	US-09-070-927A-520	Sequence 520, App
C 43	46.2	5.5	6208	12	US-10-311-455-167	Sequence 167, App
C 44	46.2	5.5	9810	12	US-10-311-455-399	Sequence 399, App
C 45	46.2	5.5	19659	12	US-10-311-455-739	Sequence 739, App

ALIGNMENTS

RESULT 1
US-09-835-992A-22
Sequence 22, Appli
Patent No. US20020037541A1
GENERAL INFORMATION:
APPLICANT: Obata, Yuichi
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED WITH GASTRIC CANCER AN
FILE REFERENCE: L0461/7112
CURRENT APPLICATION NUMBER: US/09/835, 992A
PRIOR FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: US 08/896, 164
NUMBER OF SEQ ID NOS: 87
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 843
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Unsure
LOCATION: (12)..(12)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (19)..(19)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (33)..(33)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (35)..(35)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (42)..(42)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (47)..(47)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure

LOCATION: (56)..(56)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (64)..(65)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (69)..(69)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (71)..(71)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (74)..(74)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (78)..(78)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (80)..(80)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (84)..(84)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (92)..(93)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (95)..(95)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (107)..(107)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (122)..(122)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (130)..(130)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (139)..(139)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (148)..(148)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (157)..(157)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (164)..(164)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (169)..(169)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (174)..(174)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (176)..(176)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (179)..(179)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (185)..(185)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (189)..(189)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (191)..(191)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (200)..(200)

OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (203)..(203)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (208)..(208)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (213)..(214)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (216)..(216)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (225)..(225)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (229)..(229)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (246)..(246)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (249)..(249)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (255)..(256)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (260)..(260)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (262)..(263)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (270)..(270)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (274)..(274)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (287)..(287)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (292)..(292)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (296)..(298)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (305)..(305)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (307)..(307)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (312)..(312)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (329)..(329)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (332)..(332)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (342)..(342)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (348)..(348)
OTHER INFORMATION: n = a, c, g or t
NAME/KEY: Unsure
LOCATION: (357)..(357)
OTHER INFORMATION: n = a, c, g or t

```
/ NAME/KEY: Unsure
/ LOCATION: (379)..(379)
/ OTHER INFORMATION: n = a, c, g or t
/ NAME/KEY: Unsure
/ LOCATION: (381)..(381)
/ OTHER INFORMATION: n = a, c, g or t
/ NAME/KEY: Unsure
/ LOCATION: (387)..(388)
/ OTHER INFORMATION: n = a, c, g or t
/ NAME/KEY: Unsure
/ LOCATION: (394)..(395)
/ OTHER INFORMATION: n = a, c, g or t
/ NAME/KEY: Unsure
/ LOCATION: (399)..(400)
/ OTHER INFORMATION: n = a, c, g or t
/ NAME/KEY: Unsure
/ LOCATION: (406)..(406)
```

Query Match 86.0%; Score 725; DB 9; Length 843;
Best Local Similarity 100.0%; Pred. No. 2.7e-165;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGCCAAAAAANTTATTNAAATTTCTATTAAANCNTCTCCNCAAAANCATTATTNACCC 60
   |||||
Db 1 GGCCAAAAAANTTATTNAAATTTCTATTAAANCNTCTCCNCAAAANCATTATTNACCC 60

QY 61 TATNNCNCNGANTTTNANAANTACCTTNTNTTTAAAAAACCTNGGAAAAAAATAA 120
   |||||
Db 61 TATNNCNCNGANTTTNANAANTACCTTNTNTTTAAAAAACCTNGGAAAAAAATAA 120

QY 121 TNGCAATANTTAACCTTCTGAAAAANGAAATTTTACCAANGACNGAAANCNTTNT 180
   |||||
Db 121 TNGCAATANTTAACCTTCTGAAAAANGAAATTTTACCAANGACNGAAANCNTTNT 180

QY 181 AATTNGAANTNAAATTATANTNGAANGCGCNCNGAAGCAACCAANTTNATGTCGAATT 240
   |||||
Db 181 AATTNGAANTNAAATTATANTNGAANGCGCNCNGAAGCAACCAANTTNATGTCGAATT 240

QY 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANGCCNNNTT 300
   |||||
Db 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANGCCNNNTT 300

QY 301 AAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCCTTNAAT 360
   |||||
Db 301 AAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCCTTNAAT 360

QY 361 TGGATTGGGAAAAAATNCGTTCNNATACCCNNGAANGCAAAANTTTTAAATTTTA 420
   |||||
Db 361 TGGATTGGGAAAAAATNCGTTCNNATACCCNNGAANGCAAAANTTTTAAATTTTA 420

QY 421 ACCCCCTANTTTTAAANCTATNGAANAANTNGATTANNGACTTGAATGCCAACCTTAN 480
   |||||
Db 421 ACCCCCTANTTTTAAANCTATNGAANAANTNGATTANNGACTTGAATGCCAACCTTAN 480

QY 481 TTNCNGGCCACCGTGGGCTNGTNTTCTTACTTANTCCCCCAAGAAANNCTTAAN 540
   |||||
Db 481 TTNCNGGCCACCGTGGGCTNGTNTTCTTACTTANTCCCCCAAGAAANNCTTAAN 540

QY 541 CNGAANCTCCNCCAAATAAACCTTAANTATCCTTGGTAAACCAAAACCTTTTNG 600
   |||||
Db 541 CNGAANCTCCNCCAAATAAACCTTAANTATCCTTGGTAAACCAAAACCTTTTNG 600

QY 601 TTTACNTANTCCTTGGGATTTAACGGGTCCCAATTNATCNGAACCCANTTTTCCCCC 660
   |||||
Db 601 TTTACNTANTCCTTGGGATTTAACGGGTCCCAATTNATCNGAACCCANTTTTCCCCC 660

QY 661 NAAACATANTTACCATTTTACCTTGGTAAAGGCNCAAGTNGTGCANTNCCGCAANCACT 720
   |||||
Db 661 NAAACATANTTACCATTTTACCTTGGTAAAGGCNCAAGTNGTGCANTNCCGCAANCACT 720

QY 721 ANTNTTCCCGCNCCTTCCCGGANGCCTTGGAAGAAACGGGATNGTCCCGCCTTAA 780
   |||||
Db 721 ANTNTTCCCGCNCCTTCCCGGANGCCTTGGAAGAAACGGGATNGTCCCGCCTTAA 780
```

```
QY 781 AAAACAACCTTCCCCCNCCTTTGGCCAGGNNTTNTTCCCGTCTAAATCCGAACATAATA 840
   |||||
Db 781 AAAACAACCTTCCCCCNCCTTTGGCCAGGNNTTNTTCCCGTCTAAATCCGAACATAATA 840

QY 841 AAG 843
   |||
Db 841 AAG 843
```

RESULT 2
US-09-880-107-3296/c
Sequence 3296, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3296
LENGTH: 1052
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 U11313
US-09-880-107-3296

Query Match 59.9%; Score 505; DB 10; Length 1052;
Best Local Similarity 71.8%; Pred. No. 4.9e-112;
Matches 610; Conservative 0; Mismatches 233; Indels 6; Gaps 3;

```
QY 1 GGCCAAAAAANTTATTNAAATTTCTATTAAANCNTCTCCNCAAAANCATTATTNACCC 60
   |||||
Db 1028 GGCCAGAAAAAGTATTATTTCTATTAACATCTCTCAAGCATATTATTATCC 969

QY 61 TATNNCNCNGANTTTNANAANTACCTTNTNTTTAAAAACCTNGGAAAAAAATAA 120
   |||||
Db 968 TATATCTCACTGAATTTTAAAGAAATAACATAGATTAAGAAAACTAGGAAAAAGATAA 909

QY 121 TNGCAATANTTAACCTTCTGAAAAANGAAATTTTACCAANGACNGAAANCNTTNT 180
   |||||
Db 908 ATGCAGATTAATTAACCTTACATGAAAAAGAAAAATTATAACAAAGGACTGAGAACGTTAT 849

QY 181 AATTNGAANTNAAATTATANTNGAANGCGCNCNGAAGCAACCAANTTNATGTCGAATT 240
   |||||
Db 848 AATTGAATGAGATTATATTTGAAAACTGACTGAAAGCAAACTTATTTGTTCAATT 789

QY 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANGCCNNNTT 300
   |||||
Db 788 ATTCTTAATGATGGTGTATTATGACTAATACATGATTTTCAATAAGGAAACCATGTT 729

QY 301 AAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGATCANATTCCTTNAAT 360
   |||||
Db 728 AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTGTGATCATATTTCTTTTAT 669

QY 361 TGGATTGGGAAAAAATNCGTTCNNATACCCNNGAANGCAAAANTTTTAAATTTTA 420
   |||||
Db 361 TGGATTGGGAAAAAATNCGTTCNNATACCCNNGAANGCAAAANTTTTAAATTTTA 420

QY 421 ACCCCCTANTTTTAAANCTATNGAANAANTNGATTANNGACTTGAATTC-CAACCTTA 479
   |||||
Db 608 ATCTCACTAATTTTAAAGAACTATGAGAAATTTGATTAATGACATGAGAGTGCACAACTA 549
```


LOCATION: (541)..(541)
OTHER INFORMATION: n = a, c, g or t
us-09-835-992a-20

Query Match 48.8%; Score 411.8; DB 9; Length 687;
Best Local Similarity 71.8%; Pred. No. 1.4e-89;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

QY 15 ATTTNAATTTCTATTANCTCTCCNCAANCATTATTNACCCTATNNCNCNGAN 74
Db 1 ATTTTAATTTCTATTAAACATTCTTCTCAAGCATTTATCTATATCTCACTGAA 60
QY 75 TTTNANAANTACCTTTNNNTTAAACCTNGAAAAAATAATNGCAATANTTAA 134
Db 61 TTTTANAATAACATTAGATTAGAAAACTAGAAAAAGATNAATGAGATAATTAA 120
QY 135 CCTNCTGAAAAANGAAATTNTTACCAANGACNGAANGNTNTTATTTGAANTNAA 194
Db 121 ACTTACATGAAAAAGAAATTAACAAGAGACTGAGACGTATATAATTGAAATGAGA 180
QY 195 TTATANTTNGAANGCGCNGNCGAACAACCAANTTNATGTCATATATCCTNAANGAGG 254
Db 181 TTATAATTTGAAAACTGCATCTGAAGAACAATTATGTTCAATTATCTTAATGATGG 240
QY 255 NNTTTNANNACTAATNCCNGATTTTCCATANGAANCCNNNTTAAANTNTTNAT 314
Db 241 TGTTTATGACTAATACACTGATTTTTCATTAAGGAAACCATGTTAAAAATATTTAT 300
QY 315 TTTAAAAATAACCCNGTNTCCAACCCNGATCANATTCCTTNAATTGGATTGGGAAA 374
Db 301 TTTAAAAATAAGCTGTGTTCAAGCTGTGATCATATTTCTTTATTTGATTGGGAGA 360
QY 375 AAATNCNGTCCNNATACNNGAANGCAAAANTTTTAAATTTTAAACCCCTANTTTT 434
Db 361 AAATACTGTTCTGATAGCAATGAATGCAAAATTTTGAATTTTAACTCNCATAATT 420
QY 435 AAAANCTATNGAAAAANTGATTANNGACTTGAATGC-CAACCTANTTNCNGCCACN 493
Db 421 AAGACTATGAGAAATGATTAATGACATGAGTGCACACACTAATTACTGGCCAGCT 480
QY 494 GTGGCNGTNGTTCCTTACTTANTCCCCCAAGGAAANNCTTAANGNANGCTCNC 553
Db 481 GTTGCATGTGTTCTTACTTACTGTTCTCCCAAGGAAACCTTAACTGAATCTTACG 540
QY 554 AAAATAACCTTAANTATCTTGGTAAACCAACAAACCTTTTNGTTTACNTANTCT 613
Db 541 NGATAACCTTAATATATCTTGTAGCCAAAC-AAAACCTTTTGTGTACATAGTTCT 598
QY 614 TGGGATTTAAGCGGTCCCAATTNNATCCNGAACCCANTTTCCCCNAAACCATANTTAC 673
Db 599 TTGGATTTTACTGTTCTTAATTATTTCTGAACCTCAATTTTCCCGACCATTAATTAC 658
QY 674 CAT-TTTACCTGTAGGNCAGTNGTT 701
Db 659 CCTATTAACTTGTATGACAGTGTGT 687

RESULT 5
US-09-796-692-7228
Sequence 7228, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7228
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-7228

Query Match 37.8%; Score 318.8; DB 10; Length 506;
Best Local Similarity 73.4%; Pred. No. 3.6e-67;
Matches 358; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

QY 1 GGGCAAAAAANTTATTNAATTTCTATTTANCTCTCCNCAANCATTATTNACCC 60
Db 19 GGGCAGAAAAAGTTATTTTAAATTTCTATTTAAACATCTTCTCAAGCATTTATTTATCC 78
QY 61 TATNCCNCGAANTTNANAANTACCTTTNNNTTAAAAAACTNGAAAAAAATAA 120
Db 79 TATATCTCACTGAATTTTAAAGAAATACATTAGTATTAGAAAACTAGAAAAAGATAA 138
QY 121 TNGCAATANTTAACCTTCTTGAANGAAATTTNTTACCAANGACNGAANGNTNT 180
Db 139 ATGCAGATAATTAACCTTACATGAAAAAGAAATTAACAAGAGACTGAGACGTTAT 198
QY 181 AATTNGAANTNAATTAANTTNGAANGCGCNGNCGAACAACCAANTTNATGTCCAATT 240
Db 199 AATTTGAATGAGATTATAATTGAAAACTGCATCTGAAGCAAACTTTATGTTCAATT 258
QY 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANGCCNNNTT 300
Db 259 ATTCTTAATGATGTGTTTATGACTAATACACTGATTTTCAATAGGAACCACTGTT 318
QY 301 AAAANTTTTNTATTTTAAAAATAACCCNGTNTCCAACCCNGATCANATTCCTTNAAT 360
Db 319 AAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTAAT 378
QY 361 TGGATTGGGAAAAAATNCGTTCNNATACCNNGAANGCAAAANTTTTAAATTTTAA 420
Db 379 TTGATTTGGAGAAATACTGTTCTGATAGCATGAATGCAAAATTTTATAGATTTTAA 438
QY 421 ACCCCCTANTTTTAAANCTATNGAANGAANTNGATTANNGACTTGAATGC-CAACCTA 479
Db 439 ATCTCACTAATTTTAAAGACTATTGAGAAATTTGATTAATGACATGAGTGCACACACTA 498
QY 480 NTTNCGG 487
Db 499 ATTTACTGG 506

RESULT 6
US-10-040-862-7228
Sequence 7228, Application US/10040862

```
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for windows Version 3.0
; SEQ ID NO 7228
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-040-862-7228

Query Match      37.8%; Score 318.8; DB 14; Length 506;
Best Local Similarity 73.4%; Pred. No. 3.6e-67;
Matches 358; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

QY 1 GGCAGAAAAAANTTATTNNAATTCCTATTAAACNTCCCNCAANCAATTTATTTNACC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19 GGCAGAAAAAAGTATTATTCTATTAAACATCTTCTCAAGCATTATTATTTCC 78

QY 61 TATNNCNCNGANTTTNANAANTACCTTTNNTTTAAAAAACCCTNGAAAAAAATAA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 TATATCTCACTGAATTTAAGAAATAACATTAGTATTAGAAAAAAGAGAAAAAGATAA 138

QY 121 TNGCAATANTTAACCTTNCCTGAAAAANGAAATTTNTACCAANGAACNGAANCNTTNT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 ATGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAAAAGACTGAGAACGTTAT 198

QY 181 AATTNGAANTNAAATATANTTNGAANAACGGCNCNCGAAACCAANCTTNATGTGCCAATT 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 AAAATGAAATGAGATTATATTTGAAAACTGCATCTGAAGCAAACTTTATGTGTCATTT 258

QY 241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTT 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 ATTCTTAATGATGTGTTTATGACTAATACACTGATTTTCAATAAGAAAAAGCAATGTT 318

QY 301 AAAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAATT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 AAAAATATTTTATTTTAAAAAATAAGCCTGTGTCAAGCTCTGATCATATTTCTTTAATT 378
```

```
QY 361 TGGATTGGGAAAAAATNCNGTCCNMNATACCNNGAANNCCAAANTTTTAAATTTTAA 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 TTGATTTGGGAAGAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTAA 438

QY 421 ACCCCCTANTTTTAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCTTA 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 ATCTCACTAATTTTAAAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACCTA 498

QY 480 NTTCNNGG 487
    |||||
Db 499 ATTAAGTG 506

RESULT 7
US-09-998-598-578
; Sequence 578, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 578
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-998-598-578

Query Match      37.2%; Score 313.6; DB 10; Length 486;
Best Local Similarity 72.8%; Pred. No. 6.3e-66;
Matches 354; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 3 CCAAAAAAANTTATTNNAATTTCTATTAAACNTCCCNCAANCAATTTATTTNACCCTA 62
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CCAGAAAAAGTATTATTATTTCTATTAAATATCTTCTCAAGCATTATTTATCCCTA 60

QY 63 TNNCNCNCNGANTTTNANAANTACCTTTNNTTTAAAAAACCCTNGAAAAAAATAATN 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TATCTCACTGAATTTTAAAGAAATAACATTAGTATTAGAAAAAAGTGAAGAAAGATAAAT 120

QY 123 GCAATANTTAACCTTNCCTTGAANAANGAAATTTNTACCAANGAACNGAANCNTTNTAA 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GCAGATAATTAACCTTACATGAAAAAGAAAAATTAACAAGAAAGACTGAGAACGTTATAA 180

QY 183 TTNGAANTNAAATTAANTTNGAANAACGGCNCNCGAAACCAANCTTNATGTGCCAATTAT 242
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ATTGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACTTTATGTGTTCAATTAT 240

QY 243 CCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTTAA 302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TCTTAATGATGTGTTTATGACTAATACACTGTTTTCATATAAGAAACCATGTTAA 300

QY 303 AANTNTTTNATTTTAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTNAATTG 362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AAATATTTTATTTTAAAAATAAGCCTGTGTGTCAGAGCTCTGATCATATTTCTTTAATTT 360

QY 363 GATTGGGAAAAAATNCNGTTCNNATACCNNGAANNCCAAANTTTTAAATTTTAAAC 422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GATTGGGAGAAATACTGTTTCTGATAGCATGAAATGCAAAATTTTGAATTTTAAAT 420

QY 423 CCCCTANTTTTAAANCTATNGAAAAANTNGATTANNGACTTGAATTGC-CAACCTTANT 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 CTCACATAATTTTAAAGAACTATTGAGAAATGATTAATGACATGAAGTGCAACAACCTAAT 480

QY 482 TNCNGG 487
```

Db 481 TACTGG 486

RESULT 8
US-10-102-524-161/c

; Sequence 161, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-161

Query Match 21.6%; Score 182.2; DB 14; Length 510;
Best Local Similarity 69.5%; Pred. No. 3.9e-34;
Matches 248; Conservative 0; Mismatches 104; Indels 5; Gaps 2;

QY 492 CNGTGGCGNTGNTTCCCTTACTTANTCCCCCAAGAAANNCCCTTAANCNGAANTCCN 551
Db 510 CTGTTGGCATTGTGTTCTTACTTACTGTTCTCCCAAGAAACTCTTAATTGAATCTTCA 451
QY 552 CCAAAATAACCCCTTAANTATCTTGTGTAACCAAAACAAACCTTTTNGTTTACTANTC 611
Db 450 GCAGATAATCCTTAATATCTTGTGTAAGCAAAACAAAGCTTTTGTGTTACTAGTT 391
QY 612 CTGGGATTTAACGGGTCCTCCCAATTNATCCNGAACCAATTTTCCCCNAACCAATANTT 671
Db 390 CTTGGATTTTACTGTCTCTTAATTTTATTCTGAACCTCAATTTTACCCCAAGACCAATTA 331
QY 672 ACCATTTTACCTTGTGTAAGCNCAGTNGTTGCANTCCGAANACAGTANTNTCC--C 729
Db 330 ACCATATTAACCTTGTGTAATGACAGTGTATGCAATTCGCAAGACAGTAGTATACATC 271
QY 730 CNGGCNCTTCCCGGANCCTTGGGAAAAACGGGATNGTCCCCCCTT--AAAAACA 786
Db 270 AGGCTCTATTCAACCCAGACATAGGAAAAACAGGATGGTCAACCCATTAAGAAACAA 211
QY 787 ACCTTCCCCCNCCTTTGGCCCGAGNNNTNTTCCCGCTTAATCCGAACATAAAAAAG 843
Db 210 TCTATCCACGCTATTGTGCAATGTGTAGTTTCAGTCTAATTTCTGACAAATAAATG 154

RESULT 9
US-10-102-524-188/c

; Sequence 188, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 188
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-188

Query Match 21.6%; Score 182.2; DB 14; Length 510;
Best Local Similarity 69.5%; Pred. No. 3.9e-34;
Matches 248; Conservative 0; Mismatches 104; Indels 5; Gaps 2;

QY 492 CNGTGGCGNTGNTTCCCTTACTTANTCCCCCAAGAAANNCCCTTAANCNGAANTCCN 551
Db 510 CTGTTGGCATTGTGTTCTTACTTACTGTTCTCCCAAGAAACTCTTAATTGAATCTTCA 451
QY 552 CCAAAATAACCCCTTAANTATCCTTGTGTAACCAAAACAAACCTTTTNGTTTACTANTC 611
Db 450 GCAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTTGTGTTACATAGTT 391
QY 612 CTGGGATTTAACGGGTCCTCCCAATTNATCCNGAACCAANTTTTCCCCNAACCATANTT 671
Db 390 CTTGGATTTTACTGTCTCTAATTTTATCTGAACCTCAATTTTACCCCAAGACCATATT 331
QY 672 ACCATTTTACCTTGTGTAAGCNCAGTNGTTGCANTNCCGAANACAGTANTNTCC--C 729
Db 330 ACCATATTAACCTTGTGTAATGACAGTGTATGCAATTCGCAAGACAGTAGTATACATC 271
QY 730 CNGGCNCTTCCCGGANCCTTGGGAAAAACGGGATNGTCCCCCCTT--AAAAACA 786
Db 270 AGGCTCTATTCAACCCAGACATAGGAAAAACAGGATGGTGCACACCCATTAAGAAACA 211
QY 787 ACCTTCCCCCNCCTTTGGCCCGAGNNNTNTTCCCGCTTAATCCGAACATAAAAAAG 843
Db 210 TCTATCCACGCTATTGTGCAATGTGTAGTTTCAGTCTAATTTCTGACAAATAAATG 154

RESULT 10
US-10-102-524-645

; Sequence 645, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10/102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 645
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-102-524-645

Query Match 21.6%; Score 182.2; DB 14; Length 510;
Best Local Similarity 69.5%; Pred. No. 3.9e-34;
Matches 248; Conservative 0; Mismatches 104; Indels 5; Gaps 2;

QY 492 CNGTGGCGNTGNTTCCCTTACTTANTCCCCCAAGAAANNCCCTTAANCNGAANTCCN 551
Db 1 CTGTTGGCATTGTGTTCTTACTTACTGTTCTCCCAAGAAACTCTTAATTGAATCTTCA 60
QY 552 CCAAAATAACCCCTTAANTATCCTTGTGTAACCAAAACAAACCTTTTNGTTTACTANTC 611
Db 61 GCAGATAATCCTTAATATACTTTGTAAGCAAAACAAAGCTTTTGTGTTACATAGTT 120
QY 612 CTGGGATTTAACGGGTCCTCCCAATTNATCCNGAACCAANTTTTCCCCNAACCATANTT 671
Db 121 CTTGGATTTTACTGTCTCTAATTTTATTCTGAACCTCAATTTTACCCCAAGACCATATT 180


```
OY 613 TTGGGATTTAACGGGTCGCCAATTNNATCCNGAACCCANTTTCCCCNACCATANTTA 672
    ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 339 TT-GTATTCTGCTATTATTAATACTATCTTCTAAAACTCACTTTTACCACCAATGATTA 281
OY 673 CCATTTTACCTTGTGAAGCAGTNGTTGCANTNCCGAAANCAAGTANTNTTC-CCN 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 CCATATTAACTGTGTAATGCACAGATGTAGCAACTCAGCAATGCTGTAGTAAACCATCA 221
OY 732 GGCNCTTTCCCGGANCCTTGGGAAAAACGGGATNGTCCCCCTTAAAAAACACCTT 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 GGTATTATTACCCAGCAATTGGATAAAACAGACTTGATCAACCCATTAAACCAAC 161
OY 792 CCCCCNCTTTGGCCCGAGGNNNTTCCCGTCTAAATCCGAACAATAAAA 841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 160 CTCATGCTATTCCCAATCTAATAGTTTCAATCTAATCTGACCAATGAAA 111
```

RESULT 14

```
US-10-311-455-914/c
; Sequence 914, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 914
; LENGTH: 18133
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-914
```

Query Match 7.3%; Score 61.8; DB 12; Length 18133;

Best Local Similarity 40.1%; Pred. No. 0.00036;

Matches 174; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```
OY 5 AAAAAAANTTATTNAATTTCTTAATTAAACNTCTCCNCAAAACATTATTNNACCTATN 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16239 AAAAAAATAACATACCTAACAATATAAAACGTCCAAAATAAATAATTTTAAACATTTT 16180
OY 65 NCNCNCNANTTTNANAAANTACCTTNTNTTAAAAAACTNGAAAAAAAATAATNGC 124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16179 CACTCTATAACAATAATCATTTCAATAATATCATATATAATTAATAATAATTTT 16120
OY 125 AAATANTTAACCTTNCCTGAAGAANGAAATTTNTACCAANGAANGAANCNTTNTAAT 184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16119 CTTACTTTTATCTATAATAAAAAACAATATTTTAAAAAAATACATTTCCACTTTAATA 16060
OY 185 NGAAANTNAATTATANTTNGAAANGCGCNCNGAAGCAACCAANTTNATGTCCAAATATCC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16059 TAAATATAAAATTCATAATAACCAATTCATATATAAACCAAAATTTTCACTTACCAAAAA 16000
OY 245 TNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAAATANGGAANCCNNNTTAA 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15999 AATCTTACATCTTACAAATTAATTAATTTACTATTTTCCATTTTAAATACAAATAAAA 15940
OY 305 NTNTTTTNAATTTAAAAATAACCCNGTNTCCAACCCCGCATCANATTCCTTNAATTTGA 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15939 TCCAAAATATACACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15880
```

```
OY 365 TTGGGAAAAAATATNCGTTCNNATACCNNGAANNCAAAANTTTTAAATTTTAAACC 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15879 ATCATCTAATCAATTAATTTATTTTCAAAATCACTAATTAATAAATAAATAAATAA 15820
OY 425 CCCTANTTTTAAAA 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15819 TACAAATTTTAAAA 15806
```

RESULT 15

```
US-10-311-455-1922/c
; Sequence 1922, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1922
; LENGTH: 18218
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1922
```

Query Match 6.1%; Score 51.8; DB 12; Length 18218;

Best Local Similarity 42.0%; Pred. No. 0.094;

Matches 116; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

```
OY 44 AAACATTATTTNACCTATNCCNCGANTTTNANAAANTACCTTNTNTTTAAAAA 103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4063 AAAAAATTAATTAATAAAACCTTAAATAATATCTTAAACCAAAAAATATTAATAAATAA 4004
OY 104 CCTNGAAAAAATAATNGCAATATANTTAACCTTNCCTGAAAAANGAAATTTNTACCA 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4003 CAATAAAAATAAAAAACTAAAAAAATTAATAATATATTTTAAAAATAAAAACCATTA 3944
OY 164 NGGACNGAAANCNTTAAATTNGAANTNAATTAATTAATTNGAANGCGCNCNGAACA 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3943 ATTTCTTACATATAATTAATTTAAAAACATATAATATTTTAAATATATACCTTTAAT 3884
OY 224 ANCTTNATGTCATATATCTTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCA 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3883 ATAATATACATCAATTTTATAAAAAAATAACATATTTTAATTTAATTAATAAATTTAT 3824
OY 284 ATANGGAANCCNNNTTAAANTTTTNAATTTTAA 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3823 TTTTAAATTCCTTAATAAAACCTTTTCTTAA 3788
```

Search completed: November 27, 2003, 12:53:10
Job time : 320.882 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: November 27, 2003, 06:00:22 ; Search time 62.2227 Seconds
(without alignments)
5979.910 Million cell updates/sec

Title: US-09-835-992A-22

Perfect score: 843

Sequence: 1 ggccaaaaaanttatcttca.....taatccgaacaataaaag 843

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	725	86.0	843	3	US-08-896-164-22 Sequence 22, Appl
2	436.2	51.7	714	3	US-08-896-164-19 Sequence 19, Appl
3	411.8	48.8	687	3	US-08-896-164-20 Sequence 20, Appl
C 4	110	13.0	173	4	US-09-313-294A-1553 Sequence 1553, Ap
5	48.8	5.8	7218	1	US-08-232-463-14 Sequence 14, Appl
6	44.6	5.3	29793	4	US-09-302-812-38 Sequence 38, Appl
7	44.6	5.3	29793	4	US-09-511-477-38 Sequence 38, Appl
8	44.6	5.3	29793	4	US-09-511-507-38 Sequence 38, Appl
C 9	41.6	4.9	1664976	4	US-08-916-421B-1 Sequence 1, Appl
C 10	40.8	4.8	168575	4	US-09-426-290-1 Sequence 1, Appl
C 11	39.8	4.7	6124	4	US-08-213-419B-3 Sequence 3, Appl
C 12	39.4	4.7	5852	1	US-07-867-106-2 Sequence 2, Appl
C 13	39	4.6	8920	2	US-08-446-855A-1 Sequence 1, Appl
C 14	39	4.6	8920	3	US-09-150-741-1 Sequence 1, Appl
15	38.6	4.6	168575	4	US-09-426-290-1 Sequence 1, Appl
16	37.8	4.5	1946	4	US-09-886-319A-83 Sequence 83, Appl
17	37.4	4.4	615	3	US-08-998-416-186 Sequence 186, App
18	37	4.4	828	5	PCT-US94-06079-27 Sequence 27, Appl
19	36.6	4.3	2862	4	US-09-443-184-36 Sequence 36, Appl
20	36.4	4.3	20674	4	US-09-641-638-651 Sequence 651, App
21	36.2	4.3	6243	2	US-09-056-075-1 Sequence 1, Appl
22	36	4.3	1554	2	US-08-031-538-8 Sequence 8, Appl
23	36	4.3	1554	4	US-09-016-434-1301 Sequence 1301, Ap
24	36	4.3	1730	1	US-07-817-920-1 Sequence 1, Appl
25	36	4.3	1730	1	US-08-117-006-1 Sequence 1, Appl
26	36	4.3	1730	1	US-08-216-594-1 Sequence 1, Appl
27	36	4.3	1730	5	PCT-US93-00149-1 Sequence 1, Appl

C 28	36	4.3	58073	4	US-08-545-528D-1	Sequence 1, Appli
C 29	35.8	4.2	837	3	US-08-998-416-288	Sequence 288, App
C 30	35.8	4.2	5183	2	US-08-870-518-7	Sequence 7, Appli
C 31	35.8	4.2	7244	3	US-08-378-313-26	Sequence 26, Appli
C 32	35.8	4.2	20674	4	US-08-641-638-651	Sequence 651, App
C 33	35.6	4.2	209	4	US-09-702-705-929	Sequence 929, App
C 34	35.6	4.2	209	4	US-09-736-457-929	Sequence 929, App
C 35	35.6	4.2	1044	4	US-09-495-406-12	Sequence 12, Appl
C 36	35.6	4.2	1866	4	US-09-601-198-153	Sequence 153, App
C 37	35.6	4.2	1978	1	US-08-416-870C-1	Sequence 1, Appli
C 38	35.6	4.2	11474	4	US-09-495-406-1	Sequence 1, Appli
C 39	35.6	4.2	22067	4	US-09-820-001-3	Sequence 3, Appli
C 40	35.6	4.2	87350	3	US-08-781-891-79	Sequence 79, Appl
C 41	35.6	4.2	87350	4	US-09-618-166-79	Sequence 79, Appl
C 42	35.6	4.2	87543	4	US-09-791-211-3	Sequence 3, Appli
C 43	35.4	4.2	1945	2	US-08-525-742-5	Sequence 5, Appli
C 44	35.4	4.2	2062	4	US-09-620-312D-170	Sequence 170, App
C 45	35.2	4.2	19124	2	US-08-487-826B-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-896-164-22
; Sequence 22, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; NUMBER OF INVENTIONS: DIAGNOSING AND TREATING GASTRIC CANCER
; CORRESPONDENCE ADDRESSES: 87
; ADDRESSES: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 843 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-896-164-22

Query Match 86.0%; Score 725; DB 3; Length 843;
Best Local Similarity 100.0%; Pred. No. 1.4e-186;
Matches 843; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCCAAAAAANTTATTNAATTCTCTATTAAACNTCTCCNCAANCATTATTNAACCC 60
Db 1 GGCCAAAAAANTTATTNAATTCTCTATTAAACNTCTCCNCAANCATTATTNAACCC 60
OY 61 TATNNCNCNCNGANTTTNANAANTACCTTTNNTNTTAATAAAACCTNGAAAAAATAA 120

Db	61	TATNCCNCNCGANTTTNAAAANTACCTTTNNNTTTAAAAAACCTNGGAAAAAAATAA	120
QY	121	TNGCAATANTTTAACTTNCCTGAAAAANGAAATTTNTACCAANGACNGAAANCNTNT	180
Db	121	TNGCAATANTTTAACTTNCCTGAAAAANGAAATTTNTACCAANGACNGAAANCNTNT	180
QY	181	AAATTNGAANTNAAATTATANTTTNGAAANCGGCNNCGAAACCAANTCTNATGTCCAATT	240
Db	181	AAATTNGAANTNAAATTATANTTTNGAAANCGGCNNCGAAACCAANTCTNATGTCCAATT	240
QY	241	ATCCTNAANGAGGNNTTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTT	300
Db	241	ATCCTNAANGAGGNNTTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNTT	300
QY	301	AAAAANTTTTTNATTTTAAAAATAACCCGNTTCCAACCCCGNATCANATTCCTTNAAT	360
Db	301	AAAAANTTTTTNATTTTAAAAATAACCCGNTTCCAACCCCGNATCANATTCCTTNAAT	360
QY	361	TGATTGGGGAAAAAATNCNGTCCNNAATACCNNGAANNCAANTTTTAAATTTTAA	420
Db	361	TGATTGGGGAAAAAATNCNGTCCNNAATACCNNGAANNCAANTTTTAAATTTTAA	420
QY	421	ACCCCTTANTTTTAAANCTATNGAAANNTNGATTANNGACTTGAATTGCCAACCTTAN	480
Db	421	ACCCCTTANTTTTAAANCTATNGAAANNTNGATTANNGACTTGAATTGCCAACCTTAN	480
QY	481	TTNCGGCCCAACCGTGGGCGNTNGTNTCTTACTTANTCCCCCAAGAAANNCCTTAAN	540
Db	481	TTNCGGCCCAACCGTGGGCGNTNGTNTCTTACTTANTCCCCCAAGAAANNCCTTAAN	540
QY	541	CNGAANCTCCNCCAAAAATAACCCCTTAANTATCCTTGTAACCAANCAAAACCTTTTNG	600
Db	541	CNGAANCTCCNCCAAAAATAACCCCTTAANTATCCTTGTAACCAANCAAAACCTTTTNG	600
QY	601	TTTACNTANTCCTTGGGATTTAACGGGTCCCAATTNATCCNGAACCAANTTTTCCCCC	660
Db	601	TTTACNTANTCCTTGGGATTTAACGGGTCCCAATTNATCCNGAACCAANTTTTCCCCC	660
QY	661	NAACCATANTTTACCATTTTACCTTGTAAGCNCAGTNGTTGCANTNCCGCAANCACT	720
Db	661	NAACCATANTTTACCATTTTACCTTGTAAGCNCAGTNGTTGCANTNCCGCAANCACT	720
QY	721	ANTNTTCCCCNGGCNTTTCCCCCGCANCCTTGGGAAAAACGGGATNGTCCCCCTTAA	780
Db	721	ANTNTTCCCCNGGCNTTTCCCCCGCANCCTTGGGAAAAACGGGATNGTCCCCCTTAA	780
QY	781	AAAAACAACCTTCCCCCNCCTTTGGCCCAAGGNNTTNTTCCCGTCTAATCCGACAATAAA	840
Db	781	AAAAACAACCTTCCCCCNCCTTTGGCCCAAGGNNTTNTTCCCGTCTAATCCGACAATAAA	840
QY	841	AAG 843	
Db	841	AAG 843	

```

RESULT 2
US-08-896-164-19
; Sequence 19, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felde & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
;

```

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
3
4 COMPUTER: IBM PS/2
5 OPERATING SYSTEM: PC-DOS
6 SOFTWARE: Wordperfect
7
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: US/08/896,164
10 FILING DATE: July 17, 1997
11 CLASSIFICATION: 424
12
13 ATTORNEY/AGENT INFORMATION:
14 NAME: No. 6218521man D. Hanson
15 REGISTRATION NUMBER: 30,946
16 REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
17
18 TELECOMMUNICATION INFORMATION:
19 TELEPHONE: (212) 688-9200
20
21 TELEFAX: (212) 838-3884
22
23 INFORMATION FOR SEQ ID NO: 19:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 714 base pairs
26 TYPE: nucleic acid
27 STRANDEDNESS: double
28 TOPOLOGY: linear
29
30 US-08-896-164-19

```

Query Match	51.7%;	Score 436.2;	DB 3;	Length 714;
Best Local Similarity	73.1%;	Pred. No. 1.2e-108;		
Matches 516; Conservative	0;	Mismatches 187;	Indels 3;	Gaps 3;

OY		2	GCCAAAAAAANTTATTNNAAATTTCTTAATAANCNTCCTCCNCAAAANCATTATTNACCCT	61
Db		2	GCCAGAAAAAGTTATTTTAATTTTCTAFTTAAACATTCCTCTCAAAGCATTTATTATCCT	61
OY		62	ATNNCN CNCGAN TTTNANAANTA CTTTNNTNTTAAAAAACCTNGGAAAAAAAATAAT	121
Db		62	ATATCTCACTGAATTTTAAGAATAACATTAGTATTAGAAAAACTAGGAAAAAGATATAA	121
OY		122	NGCAATANTTTAACCTTNCCTGAAAAANGAAATTTNTACCAANGACNGAAANCNTNTA	181
Db		122	TGCAGATAATTAACCTTACATGAAAAAGAAAAATTATACAAGAACTGAGACGTATA	181
OY		182	ATTNGAANTNAATTTATANTNGAANA NC GCGCN CNCGA AACCA ANCTTNA TGGTCCAATTA	241
Db		182	AATGGAATGAGATTATTAATTGAAAACTGCATCTGAAGCAAACCTTATGTTCATTA	241
OY		242	TCCTNAANGAGGNNTTNANNACTAATNCCNGATTTCCAAATANGAANCCNNNTTA	301
Db		242	TNCTTAATGATGCTGTTTATGACTAATAACACTGATTTTCAAGAGGAAACCCATGTTA	301
OY		302	AAANTNTTTNATTTTAAAAATAA CCNGTNTCCA ACCCNGATCANATTCCTTNAATT	361
Db		302	AAAAATATTTTATTTTAAAAATAAGCCTGTGTTCAAGCTCTGATCATATTTCTTTATTT	361
OY		362	GGATTGGGAAAAAAAAATNCNGTTCNNNATACCNNGAANNGCCAAANTTTTTTAATTTTAA	421
Db		362	TGATTTGGGAANAAAACTACTGTTTCTGATAGCATGA AATGCAAAATTTTATGATTTTAA	421
OY		422	CCCCCCTANTTTTAAAAANCATANGA AAAANTNGATTANNGACTGAATGC - CAACCTAN	480
Db		422	TCTCACTAATTTTAANAACTATTGA GA AATTGATTATGACATGAAGTGACA CAACACTAA	481
OY		481	TTNCGGCCACCGTGGCGNTGNNTT CCTTACTTANTCCCCCAA - GGAAANNCCTTAA	539
Db		482	TTACTGGCCAGCTGTGGCATGTGTTTCTTACTTAGTCTCCCAAGGGA AA ACTTTAA	541
OY		540	NCNGAANC TCNCCAAAATAA CCCCTTAANTATCCTTG GTAACCAANCAAAA ACCTTTTN	599
Db		542	ATTGAATCTTCAGCAGATAATCTTAAATATACTTTGTAA GCAAAA CA AAA GCTTTTTT	601
OY		600	GTTTACNTA - NTCTTGGGATTTAA CGGGTCCC CAATTTNATCCNGAACCCANTTTTCCC	658
Db		602	GTTTACATAGTCTTTGGGATTTTAA CTGTTCTTAATTTTATTTGAAA CTCAATTTTACC	661
OY		659	CNAACCATANTTACCATTTA CCTTG GTAAGGCN CAGTNGTTGC	704

Db 662 CCAGACCATAATTACCATATTGTTGTTGTGACAGTGTTC 707

RESULT 3

US-08-896-164-20
; Sequence 20, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBATA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-896-164-20

Query Match 48.8%; Score 411.8; DB 3; Length 687;
Best Local Similarity 71.8%; Pred. No. 4.4e-102;
Matches 495; Conservative 0; Mismatches 190; Indels 4; Gaps 3;

QY 15 ATTNAATTTCCTATTANNCNCTCCNCAANCAATTATTNACCTATNCCNCCNCGAN 74
Db 1 ATTTTAATTTCTATTAAACATCTCTCTCAAGCATTATTATCTCTATATCTCAGTAA 60
QY 75 TTTNANAANTACCTTTNNNTTAAACAACTNGGAAAAAATAATNGCAATANTTAA 134
Db 61 TTTTANAATAATACATTAGTATTAGAAAAAAGTAGAAAAAAGATNAATGAGATAATTAA 120
QY 135 CCTTNCCTGAAAGAAATTNTTACCAANGGACNGAANCNTNTTAATTNGAANTNAAA 194
Db 121 ACTTACATGAAAAAGAAATTAACAAGGAGCTGAGAACGTTATAAATTGAAATGAGA 180
QY 195 TTATANTTNGAANGCGCNCNGAACAACCAANTTNATGGTCCAAATTATCCTNAANGAGG 254
Db 181 TTATAATTGAAAACTGCATCTGAAAGCAAACTTATGTTCATATTATCTTAATGATGG 240
QY 255 NNTTNNANACTAATNCCNGAATTTTCCATANGAANGAANCNNNTTAAANTNTTNAAT 314
Db 241 TGTTTTATGACTAATACACTGATTTTTCATAAGAAACCAATGTTAAAAATATTTTAT 300
QY 315 TTTAAAAAATAACCCNGTNTCCAAACCCNGATCANATTCCTTTNATTTGGATTGGGAAA 374
Db 301 TTTAAAAAATAAGCCTGTGTTCAGAGCTCTGATCATATTTCTTTTATTTTGGATTGGGAAA 360

QY 375 AAATNCNGTTCNNNATACCNNGAANNGCAAANTTTTAAATTTTAAACCCCTANTTTT 434
Db 361 AAATACGTGTTCTGATGACATGAATGCAAAATTTTGAATTTTAACTCNCATAATTT 420
QY 435 AAANCTATNGAAAAANTNGATTANNGACTTGAATGC-CAACCCATANTTNCNGCCACCN 493
Db 421 AAGAACTATTGAGAAATTGATTAATGACATGAAGTGACAACAATAATTACTGGCCAGCT 480
QY 494 GTGGCCTNGTNTTCCCTTACTTANTCCCCCAAGGAAANNCTTAAACNGAANTCCNCC 553
Db 481 GTTGCATTTGTTGTTCTTACTTACTTACTTCTCCCAAGGAAAACTTTAAACTGAATCTTCAGC 540
QY 554 AAATAACCCCTTAANTATCTCTGTGTAACCAANCAAAACCTTTTNGTTTACNTATCTCT 613
Db 541 NGAAATAACCTTAATATATCTTTGTTAGCCAAAC--AAAACTTTTGTGTACATAGTTCT 598
QY 614 TGGATTTTAACGGGTCCCAATTNATCCNGAACCANTTTTCCCCCAACCATANTTAC 673
Db 599 TTGATTTTACTGTCTCTAATTTTATTTCTGAAACTCCATTTTCCCAACCATATTTAC 658
QY 674 CAT-TTTACCTTGTGTAAGCNCAGTNGTT 701
Db 659 CCTATTTAATCTTGTATGACAGAGTTGTT 687

RESULT 4

US-09-313-294A-1553/C
; Sequence 1553, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 1553
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700551132H1
US-09-313-294A-1553

Query Match 13.0%; Score 110; DB 4; Length 173;
Best Local Similarity 71.3%; Pred. No. 7.3e-21;
Matches 122; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 8 AAAANTTATTNATTTCCCTATTANNCNCTCCNCAANCAATTTATTNACCTATNNCN 67
Db 172 AAAAAGTATTATTATTTCTATTAAACATCTCTTCAAGCATTATTATCTCTATATCT 113
QY 68 CNCNGANTTTNANAANATACCTTTNTNTTAAAAAAAACCCTNGGAAAAAATAATNGCAA 127
Db 112 CACTGAATTTTAAGAAATAACATTAGTATTAGAAAAAAGTAGAAAAAAGATAAATGCGAG 53
QY 128 TANTTAACCTTNCCTTGAAAAANGAAATTTNTTACCAANGAANGAANCNTT 178
Db 52 TAAATTAACCTTACATGAAAAAGAAAAATTATAACAAAGAGCTGAGAACGTT 2

RESULT 5

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.

```

/ TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
/ NUMBER OF SEQUENCES: 52
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Foley & Lardner
/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 836-9300
/ TELEFAX: (703) 683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZgpt-Fls
/
US-08-232-463-14

```

Query Match	5.8%	Score 48.8	DB 1	Length 7218
Best Local Similarity	2.1%	Pred. No. 0.0011		
Matches	8	Conservative 201	Mismatches 174	Indels 0
			Gaps	0
QY	456	TANNGACTTGAAATGGCCCAACCTTANTTNCNGGCCACCGNGTGGGCTNTGTTCCCTTACTT	515	
Db	1060	TTGCGATYYY	1119	
QY	516	ANTCCCCCAAGGAANNCCTTAANCNGAANCCTCCNCCAAATAACCCCTTAANTATCCTT	575	
Db	1120	YYY	1179	
QY	576	GATAACCAANCAAAACCTTTTNGTTTACNTANTCCTTGGGATTTAACGGTCCCAAT	635	
Db	1180	YYY	1239	
QY	636	TTNATCCNGAACCCANTTTTCCCCNAACCATANTTACCATTTTACCTTGTAAGGCNCA	695	
Db	1240	YYY	1299	
QY	696	GTNGTTGCANTNCCGCAANCAGTANTNTTCCCCNGGNCCTTCCCCGANCCTTGGA	755	
Db	1300	YYY	1359	
QY	756	AAACGGGATNGTCCCCCCCTTAATAAACCAACCTTCCCCNCCTTTGGCCCAAGNNTN	815	
Db	1360	YYY	1419	
QY	816	TTCCCGCTAAATCCGAACAATA	838	
Db	1420	YYYYYYYYYYYYYYYYGTACCAA	1442	

```

RESULT 6
US-09-302-812-38
; Sequence 38, Application US/09302812B
; Patent No. 6333148
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY (ADP-RIBOSE) GLYCOHYDROLASE (PARG) ENZ
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,768
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-302-812-38

```

Query Match	5.3%	Score 44.6;	DB 4;	Length 29793;
Best Local Similarity	40.8%;	Pred. No. 0.025;		
Matches 145; Conservative	0;	Mismatches 206;	Indels 4;	Gaps 1;

QY	5	AAAAAAAAA	NTTATTT	NAATTTCC	TATTA	ANCNTCC	CCNC	CAANCA	TATTTN	ACCCTA	TN	64		
Db	18899	AAATTG	TATTTT	AGTAAAA	ACTATA	TACGTT	TTGG	ACTTTA	AAAAAA	TTATGT	TAAA	18958		
QY	65	NCNCNC	NGANTTT	NAAAA	NTACCTT	TNNNTT	TAAAA	AAACCT	NGAAAA	AAAAAT	---AA	120		
Db	18959	ACTGAAA	AATTTAC	GTTTAT	TAGTCT	TATAT	TAAAA	ATACGG	TTTCA	AAATTA	TTTAAAA	19018		
QY	121	TNGCAA	AANTTTA	ACCCTT	NCCTG	AAAA	ANGGAA	ATTTT	NTA	CCA	ANGGAC	NGAAAA	NCNTNT	180
Db	19019	TTAAAA	TTAACT	CA	CCCTTT	GGTCA	AAATC	AGACAT	TTTGA	AAACTA	GATGT	ACTTTA	TT	19078
QY	181	AATTNG	AANTN	AAATTAT	ANTTNG	AAANGCGC	NNCNGAA	CCA	ANCT	NATG	TGCA	TT	240	
Db	19079	ACGTGA	TATATA	CTTAT	GTG	AAAA	ATGGA	AAATTTG	GAGAC	AGGTGA	ATTTAG	TTT	19138	
QY	241	ATCCTN	ANGAGG	GNNTT	TNANNA	CTAATN	CCCGA	TTTTC	CAATANG	GAANCC	NNNT	300		
Db	19139	TTTTCT	TTTTTC	GTA	CTCTA	AAAA	ATACTT	CA	TTTAT	TTTAC	ATTTTG	GAGACT	TAATTTT	19198
QY	301	AAAA	NTNTTT	TNATTT	TAAAA	ATAA	CCCGT	NTCCA	ACCC	CGATCA	NATTC	CTT	355	
Db	19199	TGACAT	GTTTT	GACAC	AAAA	AAAAAG	ATTTG	AAAA	ACCCCA	AAAAAA	ACTTAC	CTT	19253	

```

RESULT 7
US-09-511-477-38
; Sequence 38, Application US/09511477
; Patent No. 6337202
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) EN
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: N1AD 201
; CURRENT APPLICATION NUMBER: US/09/511,477
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38

```

```
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-511-477-38
```

```
Query Match      5.3%; Score 44.6; DB 4; Length 29793;
Best Local Similarity 40.8%; Pred. No. 0.025;
Matches 145; Conservative 0; Mismatches 206; Indels 4; Gaps 1;
```

```
QY      5 AAAAAAANTATTNATTTCCCTAATAANCNTCCCNCAANCAATTATTNACCCCTATN 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      18899 AAATGTATTTTAGTAAAACTAATAACGTTTGACCTAAAAAAATATGTATAA 18958
QY      65 NCNCNGANTTTNANAANTACCTTTNTNTTTAAAAAACCTNGAAAAAAAT----AA 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      18959 ACTGAAAAAATTACGTTTATGTGCTTATATTAAATACGGTTTCAAAATTAATTA 19018
QY      121 TNGCAATANTTAACCTTNCCTGAAAAANGAAATTTTACCAANGACNGAANCNTNT 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19019 TTAATAATACTCACCCTTTGGTCAAAATCAGACATTTTAGAAACTAGCATGTTA 19078
QY      181 AATNGAANTNAAATATATANTNGAANGCGGNNCNGAAACCAANCTTNATGGTCCA 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19079 ACGTTGAATATACTATGTGGAATAATGAAAAATTTGAAGACAGTGAATTTT 19138
QY      241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNT 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19139 TTTCTTTTTCGTACTCTAAAAATACTTCATTATTTCATTTTGAGACTAATTT 19198
QY      301 AAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGCATCANATTCCTT 355
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      19199 TGAACATGTTTCGAACAAAAAAGATTTTGAAAAACCCCAAAAAAACTTACTTT 19253
```

RESULT 8

```
US-09-511-507-38
; Sequence 38, Application US/09511507
; Patent No. 6395543
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) EN
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/511,507
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-511-507-38
```

```
Query Match      5.3%; Score 44.6; DB 4; Length 29793;
Best Local Similarity 40.8%; Pred. No. 0.025;
Matches 145; Conservative 0; Mismatches 206; Indels 4; Gaps 1;
```

```
QY      5 AAAAAAANTATTNATTTCCCTAATAANCNTCCCNCAANCAATTATTNACCCCTATN 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      18899 AAATGTATTTTAGTAAAACTAATAACGTTTGACCTAAAAAAATATGTATAA 18958
QY      65 NCNCNGANTTTNANAANTACCTTTNTNTTTAAAAAACCTNGAAAAAAAT----AA 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      18959 ACTTGAATAATTACGTTTATGTGCTTATATTAAATACGGTTTCAAAATTAATTA 19018
```

```
QY      121 TNGCAATANTTAACCTTNCCTGAAAAANGAAATTTTACCAANGACNGAANCNTNT 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19019 TTAATAATACTCACCCTTTGGTCAAAATCAGACATTTTAGAAACTAGCATGTACTTAT 19078
QY      181 AATNGAANTNAAATATANTNGAANGCGGNNCNGAAACCAANCTTNATGGTCCA 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19079 ACGTTGAATATACTATGTGGAATAATGAAAAATTTGAAGACAGTGAATTTT 19138
QY      241 ATCCTNAANGAGGNNTTNANNACTAATNCCNGATTTTCCAATANGAANCCNNNT 300
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      19139 TTTCTTTTTCGTACTCTAAAAATACTTCATTATTTTTACATTTTGAGACTAATTT 19198
QY      301 AAAANTNTTTNATTTTAAAAATAACCCNGTNTCCAACCCCGCATCANATTCCTT 355
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      19199 TGAACATGTTTCGAACAAAAAAGATTTTGAAAAACCCCAAAAAAACTTACTTT 19253
```

RESULT 9

```
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
```

```
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
```

```
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 4.9%; Score 41.6; DB 4; Length 1664976;
Best Local Similarity 42.6%; Pred. No. 0.79;
Matches 121; Conservative 2; Mismatches 160; Indels 1; Gaps 1;

```
QY 96 TTAATAAACCTNGGAAAAAATAATNGCAATANTTAACCTTCTGAAANGGAAAT 155
    ||||| ||| ||| ||||| ||| ||| ||||| ||| ||| |||||
Db 1150605 TTAATAAGAACTCGCTAAGCAATAGTTAGAGAAATGGACAACCAAGAAATAGAAACC 1150546

QY 156 TNTACCAANGACNGAANCNTTNTAATTNGAANTNAATATATANTNGAANGGCGNCC 215
    :||| | | | | | | | | | | | | | | | | | | | | | |
Db 1150545 TTYACAACTCTAAGTAAGAAGCACTTCAAAAAATTAACAGTTCATCTAGATTATAAT 1150486

QY 216 NGAAACCAANCTTATGCTTCAATATCTCTNAANGAGGNNTT-NANNACTAATNCCN 274
    ||| | | | | | | | | | | | | | | | | | | | | |
Db 1150485 GGACAACTAATTCAATGCTGTATATTCTGTAAGATTCACTTAGRAAACTCGAACATT 1150426

QY 275 GATTTTCCAATANGAANGCCNNNTTAAANNTTNTAATTTTAAAAATAACCGNTTC 334
    ||| ||||| ||| | | | | | | | | | | | | | | | |
Db 1150425 CAAATTCAATTAAGAAACGACAGTTAAATTTGTAAGATTACTTATCAATTAATTAAG 1150366

QY 335 CAACCCNGATCANATCTCTTNTATTTGATTTGGGAAAAAAT 378
    ||| ||| | | | | | | | | | | | | | | | | |
Db 1150365 TAAATTTAATTAATTAATGTAATTTGAATTAATAAAAT 1150322
```

```
RESULT 10
US-09-426-290-1/c
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21181)...(21403)
; NAME/KEY: CDS
; LOCATION: (95252)...(95430)
; NAME/KEY: CDS
; LOCATION: (101753)...(101996)
; NAME/KEY: CDS
; LOCATION: (110324)...(110439)
; NAME/KEY: CDS
; LOCATION: (124058)...(124278)
; NAME/KEY: CDS
; LOCATION: (127009)...(127130)
; NAME/KEY: CDS
; LOCATION: (128910)...(129139)
; US-09-426-290-1

Query Match          4.8%; Score 40.8; DB 4; Length 168575;
Best Local Similarity 42.5%; Pred. No. 0.53;
Matches 93; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 112 AAAAAATATNGCAATANTTAACCTTCTTGAAAAANGAAATTTTACCANGACNGA 171
    |||||
Db 165933 AAAAAATATGCGCAGAAAATTTCCAGTACTGGGGAAGAAAATAGAAAAGCCAGATACAGG 165874

QY 172 AANCNTTNTAATTNGAANTNAAATTATANTTNGAAANGCGCNCNGAAGAACCAANTTNAT 231
    |||
Db 165873 AATCCCAAGAATACCAAAATAAATGAATCCGATAAATCCACACCAAGTACATATACAT 165814

QY 232 GGTCCAATTATCTCTNAANGAGGNNTTNANACTAATNCCNGATTTTCCAATANGAA 291
    |||||
Db 165813 CATGAATACTGTGAAGCCATAAAAAGAACAACTCATGTCTTTTTCAGCAACATGAC 165754

QY 292 NCCCNNTTAAANNTTTTNAATTTTAAAAATAACCCNG 330
    |||||
Db 165753 AGACTACTGGAATAATATATCTATGCAAAATTAACACGG 165715

RESULT 11
US-08-213-419B-3/c
; Sequence 3, Application US/08213419B
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Inselburg, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: J11-002CNC
; CURRENT APPLICATION NUMBER: US/08/213,419B
; CURRENT FILING DATE: 1994-03-14
; PRIOR APPLICATION NUMBER: US 07/870,506
; PRIOR FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
```

```
LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407)..(2439)
; NAME/KEY: CDS
; LOCATION: (2598)..(3404)
; NAME/KEY: CDS
; LOCATION: (3580)..(3720)
; NAME/KEY: CDS
; LOCATION: (3850)..(5835)
; US-08-213-419B-3

Query Match          4.7%; Score 39.8; DB 4; Length 6124;
Best Local Similarity 39.6%; Pred. No. 0.26;
Matches 113; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 76 TTNANAANTACCTTNNNTTAAACCTNGAAAAAATAATNGCAATANTTAAC 135
    |||||
Db 6092 TGTACAACTATCGTGTATTTTAAAAAAGAAAAAAGAAATATA 6033

QY 136 CTTNCTGAAAANGAAATTTNTACCAANGACNGAANCNTTNTAATTNGAANTNAAAT 195
    |||||
Db 6032 TTTACATGAATAAAAAATTAATCTTAAAAAATAATAATATATATTTATATATA 5973

QY 196 TATANTTNGAAANGCGCNCNGAAGAACCAANTTNATGTCCTCAATTATCTTNAANGAGGN 255
    |||||
Db 5972 TTTATAGAGGTACATGATTAATTAAGATATATTACTACCTTAATAAATGAATATGA 5913

QY 256 NTTNANACTAATNCCNGATTTTCCAATANGAANCCNNTTAAANNTNTTNAT 315
    |||||
Db 5912 GAGTTATGCCCTATTTAAAAATTAAGGATACATAAAAAATATTTTNTTTT 5853

QY 316 TTAATAATAACCCNGTNTCCAACCCNGATCANATTCCTTNAT 360
    |||||
Db 5852 TTTGTATATATTATTTATACATAACAGAAATAACATTCGTTATTT 5808

RESULT 12
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostellium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526r18
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
```

```

/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 215-568-3100
/ TELEFAX: 215-568-3439
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5852 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ ANTI-SENSE: NO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 2378..5038
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 2378..5038
/
US-07-867-106-2

```

Query Match	4.7%;	Score 39.4;	DB 1;	Length 5852;
Best Local Similarity	43.4%;	Pred. No. 0.33;		
Matches	85;	Conservative	0;	Mismatches 111;
				Indels 0;
				Gaps 0;

QY	16	TTTNAATTTCCCTATTAA	NCNCTCCCNCA	AANCATTTATT	NA	CCCTATNN	CNCNC	ANGANT	75
Db	5751	TTGTATTTT	TTATATATNG	TATTTG	TGTGTTGTTT	TTTACTT	ATATTTCT	ATTTT	TTTA
QY	76	TTNNAANA	NTACCTTT	NNNTNT	TTAAAAA	ACCTNG	GA	AAAAA	ATA
Db	5691	TTTATAA	AATTA	TATTA	ATTA	ATTTTA	ATA	ATTA	AAAAA
QY	136	CTTNC	TTGA	AAANG	GA	AATTT	NT	T	N
Db	5631	ATTA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA
QY	196	TAT	ANT	T	NG	A	A	A	N
Db	5571	TAAAA	AATG	T	AAAA	GGG	5556		

RESULT 13
 US-08-446-855A-1/c
 ; Sequence 1, Application US/08446855A
 ; Patent No. 5849573
 ; GENERAL INFORMATION:
 ; APPLICANT: Stewart, Thomas S
 ; APPLICANT: Flores, Maria V
 ; APPLICANT: O'Sullivan, William J
 ; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
 ; TITLE OF INVENTION: phosphate synthetase II
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Nixon & Vanderhye PC
 ; STREET: 1100 No. 5849573th Glebe Road, 8th Floor
 ; CITY: Arlington
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/446,855A
 ; FILING DATE: 06-Jul-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mitchard, Leonard C
 ; REGISTRATION NUMBER: 29.009
 ; REFERENCE/DOCKET NUMBER: 47-80
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-816-4000

```

; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 1
; SEQUENCE CHARACTERISTICS
; LENGTH: 8920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic
;
US-08-446-855A-1

```

Query Match	4.6%;	Score 39;	DB 2;	Length 8920;
Best Local Similarity	43.1%;	Pred. No. 0.5;		
Matches	87;	Conservative	0;	Mismatches 115;
			Indels	0;
			Gaps	0;

QY	6	AAAAAANTTATTNAATTTCTATTAA	CNTCC	CNC	CAAN	CATTATT	TNACC	TATNN	65
Db	8732	AATAAATTATTAATAATTAATAAAAT	TATTTG	TAAATG	AAAAAA	CCATT	TTGG	TTATAC	8673
QY	66	CNCNCGANTTTNANAANTACCTT	TNNTNT	TAAAAA	CC	TNG	AAAAA	AAATAATNGCA	125
Db	8672	ATATGAGTGAATATAATAAATTTG	TATATAT	TAACAAAA	AAATTT	TAAAAA	AAATAATCAAT	8613	
QY	126	AATANTTAACCTTNCCTGAAAAANG	GAATTT	N	TACCA	ANGGA	CNGAA	ANCNTTNTAATTN	185
Db	8612	ATAAATTAAATGATCATATTTAA	TATAAAAA	CAATTT	TATTAATTT	TAAACG	TAAATATAATTA	8553	
QY	186	GAANTNAAATTATANTTNGAAA	207						
Db	8552	AATAAATTAATATATATATATA	8531						

```

RESULT 14
US-09-150-741-1/c
; Sequence 1, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; Patent No. 6183996
; TITLE OF INVENTION: Synthetase II
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: AU93/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8920
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
US-09-150-741-1

```

Query Match	4.6%;	Score 39;	DB 3;	Length 8920;
Best Local Similarity	43.1%;	Pred. No. 0.5;	Mismatches 0;	Gaps 0;
Matches	87;	Conservative	0;	Indels 115;
QY	6	AAAAAANTTATTNNAATTCCTATTAA	CNCTCCNCNCAANCAATTATTNACCTATTNN	65
Db	8732	AATAAATTATTAAATAATTATAAAATTATTGTAAATTGA	AAAAAACCATTTTGGTTTATAC	86733
QY	66	CNCNCNGANTTNANAAANTACCTTTNNNTNTTAA	AAAAAACCTNGGAAAAAATAATNGCA	125
Db	8672	ATATGAGTGAATATAATAAATTGTATATATATA	CACAAAATTATATAAAAAATAATCAAT	86133
QY	126	AATANTTAACCTTNCCTGAAAAANGGA	AATTNTTACCAANGGACNGAAANCNTTNTAATTN	185
Db	8612	ATAAATTAAATGATCATATTTAAT	AAAAACATTTAATTATTTTAAACGTAATAATATAATTA	85533

Qy 186 GAANTNAAATTATANTNGAAA 207
| | | | | | | |
Db 8552 AATAAATAAATATATATATATA 8531

RESULT 15

```

US-09-426-290-1
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN MARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21181)...(21403)
; NAME/KEY: CDS
; LOCATION: (95252)...(95430)
; NAME/KEY: CDS
; LOCATION: (101753)...(101996)
; NAME/KEY: CDS
; LOCATION: (110324)...(110439)
; NAME/KEY: CDS
; LOCATION: (124058)...(124278)
; NAME/KEY: CDS
; LOCATION: (127009)...(127130)
; NAME/KEY: CDS
; LOCATION: (128910)...(129139)
US-09-426-290-1

```

Query Match 4.6%; Score 38.6; DB 4; Length 168575;

[illegible]

Search completed: November 27, 2003, 12:32:32
Job time : 65.2227 secs

THIS PAGE BLANK (USPTO)